Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

ADVANCES IN VIRUS RESEARCH, VOL. 47

PRINCIPLES OF MOLECULAR ORGANIZATION, EXPRESSION, AND EVOLUTION OF CLOSTEROVIRUSES: OVER THE BARRIERS

A. A. Agranovsky

Belozersky Institute of Physico-Chemical Biology
and the Department of Virology
Moscow State University
119899 Moscow, Russia

I. Introduction

The closterovirus group (having its name from κλωστερ, Greek for "thread") combines several positive-strand RNA viruses with very flexuous filamentous particles, of which beet yellows virus (BYV) is the type virus (Bar-Joseph and Hull, 1974; Bar-Joseph and Murant, 1982). Closteroviruses are distinct from other RNA viruses of plants in some important phenomenological aspects. They have genomes of up to 20 kilobases (kb), a value comparable only to those of the animal coronaviruses and toroviruses, which have the largest RNA genomes of all positive-strand RNA viruses. The existence of such genomes having a coding capacity several times that of an average RNA virus genome (e.g., TMV) raises questions as to the trend whereby the long genomes

have evolved and the possible novel functions they have acquired. The dramatic increase in the closterovirus genome coding capacity may be linked to the distinct ecological niche they occupy. Thus, closteroviruses are the only elongated plant viruses known so far to cause phloem-limited infections in plants and to persist in their insect vectors for many hours, in contrast to only minutes. Closteroviruses also attract interest from an applied standpoint, since the diseases they cause in crops like citrus trees and sugar beet are listed among the most economically important plant viral diseases (Duffus, 1973; Falk and Duffus, 1988). Understanding the molecular mechanisms that underlie the distinct biological patterns of closteroviruses may help in developing measures against crop losses.

Details of closterovirus cytopathology, ecology, relationships with vectors, and disease control may be found in other comprehensive reviews (Bar-Joseph *et al.*, 1979; Duffus, 1973; Falk and Duffus, 1988; Lesemann, 1988; Lister and Bar-Joseph, 1981; Murant *et al.*, 1988; Tollin and Wilson, 1988). The molecular organization, evolution, and taxonomy of closteroviruses have been reviewed in considerable depth (Coffin and Coutts, 1993; Dolja *et al.*, 1994). However, the picture changes rapidly, and thus there is a need for updating. In this review, I will focus on the molecular organization, evolution, and expression of closterovirus genomes, as well as on their unique particle structure.

II. Draft of Closterovirus Taxonomy

Until recently, the taxonomic status of closteroviruses has been ill-defined, and the group was a quite heterogeneous collection, with the main distinction being between the so-called typical and atypical closteroviruses (Bar-Joseph *et al.*, 1979; Bar-Joseph and Murant, 1982; Francki *et al.*, 1991). Typical closteroviruses have particles of 1000–2000 nm in length, induce characteristic BYV-type membranaceous vesicles in infected tissues, and are transmitted by insect vectors (Table I). Atypical (or "clostero-like") viruses have shorter particles of about 750 nm, do not induce BYV-type vesicles, and have no known vector. Sequencing of the typical and atypical closteroviruses and comparison of the encoded protein sequences confirmed the disparity between them (Agranovsky *et al.*, 1991a,b, 1994a; Dolja *et al.*, 1991, 1994; German *et al.*, 1990; Koonin, 1991; Yoshikawa *et al.*, 1992). Former closteroviruses such as apple chlorotic leafspot virus, apple stem grooving virus and citrus tatter leaf virus have already been reclassified as trichoviruses and capilloviruses, respectively (Francki *et al.*, 1991;

Mayo and Martelli, 1993; Zaccomer *et al.*, 1995), and others will possibly follow. Hence, the "atypical closteroviruses" will not be considered in this review (for an updated list, see Coffin and Coutts, 1993).

A stimulating draft of the phylogenetic taxonomy of positive-strand RNA viruses has been proposed on the strength of sequence comparisons for proteins involved in genome replication and expression, particularly RNA-dependent RNA polymerases (Koonin, 1991; Koonin and Dolja, 1993). With respect to closteroviruses, it is proposed that a family *Closteroviridae* be established which, with other groups of plant viruses (tobamoviruses, tobraviruses, hordeiviruses, bromoviruses, and wheat soil-borne furovirus), has been included in a putative order called *Tobamovirales* (Dolja *et al.*, 1994, Koonin and Dolja, 1993). As recently approved by the International Committee on Taxonomy of Viruses, the family *Closteroviridae* is divided into two genera comprising monopartite (genus *Closterovirus*; type species, BYV) and bipartite closteroviruses (unnamed genus; type species, lettuce infectious yellows virus, LIYV) (Dolja *et al.*, 1994; J. P. Martelli, personal communication). This latest revision of closterovirus taxonomy is used here (Table I). More divisions at the generic level have been proposed in the family, based on differences in genome size and the number of ORFs in monopartite closteroviruses (Dolja *et al.*, 1994; Karasev *et al.*, 1995). Sequencing of more closterovirus genomes should help to define whether these differences may serve a rationale for further subdivisions.

III. Biological Patterns and Cytopathic Effects

Closteroviruses are distributed worldwide, and some of them cause devastating crop losses. The most typical symptoms in herbaceous species are yellowing, veinal necrosis, and leaf-rolling. In woody species, disease symptoms are described as seedling yellows, stem-pitting, and die-back (Bar-Joseph *et al.*, 1979; Milne, 1988). Natural host ranges reportedly vary from narrow to moderate; in artificial inoculation, however, at least one closterovirus, BYV, infects over 100 species in 15 families (Duffus, 1973). Closteroviruses are not seed-borne and are not readily transmitted mechanically, with insects being the only principal vectors. Transmission of BYV and citrus tristeza virus (CTV) by aphids takes place in a semipersistent mode, with acquisition and inoculation feeding times of 15 min to 1 hr, and retention of the virus in the insect for up to 3 days (Bar-Joseph *et al.*, 1979; Murant *et al.*, 1988). A similar transmission mode has been demonstrated for the whitefly-borne closteroviruses LIYV (Duffus *et al.*, 1986), sweet potato sunken vein virus

TABLE I

SOME PROPERTIES OF CURRENTLY RECOGNIZED AND TENTATIVE *CLOSTEROVIRIDAE* FAMILY MEMBERS[a]

Virus (abbreviation)	Particle length (nm)	Genome size (kb)	CP mol. wt (kDa)[b]	Vector[c]	BYV-type vesicles	References
Genus *Closterovirus*						
Alligator weed stunting (AWSV)[d]	1700	NK[e]	NK	NK	Yes	Hill and Zetler (1973)
Beet pseudo-yellows (BPYV)[d,f]	1500–1800	NK	28	W	Yes	Liu and Duffus (1990)
Beet yellows (BYV)	1250–1450	15.5	22, 24	A	Yes	Bar-Joseph and Hull (1974); Agranovsky *et al.* (1994a, 1995)
Beet yellow stunt (BYSV)[d]	1400	NK	NK	A	Yes	Duffus (1972)
Burdock yellows (BuYV)	1600–1750	NK	NK	A	Yes	Nakano and Inoye (1980)
Carnation necrotic flock (CNFV)	1250–1450	NK	23.5	A	Yes	Bar-Joseph *et al.* (1976); Inouye (1974)
Carrot yellow leaf (CYLV)	1600	NK	NK	A	Yes	Yamashita *et al.* (1976)
Citrus tristeza (CTV)[d]	2000	19.3	25	A	Yes	Bar-Joseph and Lee (1989); Karasev *et al.* (1995)
Dendrobium vein necrosis (DVNV)[d]	1865	NK	NK	NK	No	Lesemann (1977)
Diodia yellow vein (DYVV)[d]	NK	NK	NK	W	Yes	Larsen *et al.* (1991)
Festuca necrosis (FNV)[d]	1725	NK	NK	NK	Yes	Schmidt *et al.* (1963)
Grapevine leafroll-associated 1 (GLRaV-1)[d]	1200–2000	NK	39	NK	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990)
Grapevine leafroll-associated 2 (GLRaV-2)[d,g]	1400-2200	NK	26	NK	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990)
Grapevine leafroll-associated 3 (GLRaV-3)[d]	1400-2200	NK	43	M	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990); Belli *et al.* (1994)
Grapevine leafroll-associated 4 (GLRaV-4)[d]	1400-2200	NK	36	NK	Yes	Hu *et al.* (1990); Zimmerman *et al.* (1990)

Heracleum virus 6 (HV6)[d]	1400	NK	NK	A	NK	Bem and Murant (1979)
Little cherry (LCV)[d]	1670	~17	46	M	Yes	Ragetti *et al.* (1982); Raine *et al.* (1986); K. Eastwell, personal communication; W. Jelkmann, personal communication
Pineapple mealybug wilt-associated (PMWaV)[d]	1200	NK	24	M	NK	Gunasinghe and German (1989)
Sugarcane mild mosaic (SCMMV)[d]	1500–1600	NK	NK	M	Yes	Lockhart *et al.* (1992)
Wheat yellow leaf (WYLV)	1600–1850	NK	NK	A	Yes	Inouye (1976)
Unnamed genus (bipartite closteroviruses)						
Lettuce infectious yellows (LIYV)	1800–2000[h]	7+8.1	28	W	Yes	Duffus *et al.* (1986); Klaasen *et al.* (1995)
Sweet potato sunken vein (SPSVV)[d]	950	~7.6+8.4	29	W	Yes	Cohen et al. (1992); Winter *et al.* (1992); U. Hoyer, personal communication

[a] Some viruses currently recognized as closteroviruses, i.e., lettuce chlorosis virus (LeCV), tomato infectious chlorosis virus (TICV), cucurbit yellow stunting disorder virus (CYSDV), and grapevine leafroll-associated viruses 5 and 6 (GLRaV-5 and -6), are not included because of the lack of descriptive data; on the other hand, LCV, which is not yet considered a family member by ICTV, is included here since its features appear to be typical of the *Closteroviridae*.

[b] The exact molecular weight values are indicated wherever sequence data were available.

[c] Abbreviations used for vectors are A, aphids; M, mealybugs; W, whiteflies.

[d] Tentative species.

[e] NK, not known.

[f] BPYV has several synonymic designations, i.e., cucumber yellows, muskmelon yellows, or melon yellows virus (J. P. Martelli, personal communication); the question of whether BPYV and cucumber chlorotic spot virus (CCSV; Section V,B) are identical is uncertain and thus awaits comparative tests (L. P. Woudt, personal communication).

[g] As indicated by recent analysis (Boscia *et al.*, 1995), GLRV-2 is identical to the virus earlier described as "grapevine corky bark-associated virus" (GCBaV; Namba *et al.*, 1991).

[h] LIYV has particles of 950 nm modal length according to other measurements (D.-E. Lesemann, personal communication).

(SPSVV; Cohen *et al.*, 1992), and beet pseudo-yellows virus (BPYV; Duffus, 1973). A semipersistent transmission pattern implies a stricter virus–vector specificity as compared with, for example, the non-persistent aphid transmission of potyviruses (Falk and Duffus, 1988). Indeed, among seven aphid species that can transmit CTV, *Toxoptera citridicis* is the most efficient vector; likewise, *Myzus persicae* is the best among 24 species transmitting BYV (Duffus, 1973; Bar-Joseph *et al.*, 1979; Lister and Bar-Joseph, 1981). Interactions of the whitefly-borne closteroviruses with their vectors seem to be even more specific; thus, LIYV and SPSVV can only be transmitted by *Bemicia tabaci* (Cohen *et al.*, 1992; Duffus *et al.*, 1986), whereas BPYV can only be transmitted by *Trialeurodes vaporariorum* (Duffus, 1973). Likewise, the mealybug-borne grapevine leafroll-associated virus 3 (GLRaV-3) and the little cherry virus (LCV) are specifically transmitted by *Pulvinaria vitis* and *Phenacoccus aceris*, respectively (Belli *et al.*, 1994; Raine *et al.*, 1986).

The time needed for aphids to successfully inoculate closteroviruses and the type of the disease symptoms may both reflect the phloem-limited nature of closteroviral infections (Bar-Joseph *et al.*, 1979; Milne, 1988). Data obtained with the help of electronic monitoring indicate that, on BYV transmission, the aphid stylet reaches the phloem in 9–38 min (Limburg *et al.*, 1994).

In the cells, closteroviruses give rise to massive particle aggregates often organized as banded inclusions. The most characteristic type of intracellular inclusions are numerous vesicles surrounded by a membrane, possibly of mitochondrial origin (Essau and Hoefert, 1971; Lesemann, 1988; Kim *et al.*, 1989). These structures, referred to as BYV-type vesicles, are considered an important taxonomic feature of the family, and are suitable for closterovirus diagnosis (Table I). The BYV-type vesicles presumably contain double-stranded RNA (Coffin and Coutts, 1993); it would be interesting to determine whether these structures have anything to do with closterovirus replication.

IV. Particle Structure: Closteroviruses Are Rattlesnakes, Not Just Threads

Closterovirus particles are 950–2000 nm in length, depending on the virus (Table I), and about 12 nm in diameter (Tollin and Wilson, 1988). Virions consist of a single RNA molecule coated by capsid protein (CP) subunits of 22–28 kDa; in some grapevine leafroll-associated viruses, however, the CPs are about 40 kDa (Table I). The particles of BYV and

CTV have a primary helix pitch of 3.5–3.7 nm, with 8.5 and 10 subunits of CP per helix turn, respectively (Bar-Joseph *et al.,* 1972; Chevallier *et al.,* 1983). In the BYV particle, the structure repeats in two helical turns (Chevallier *et al.,* 1983), in contrast to the heracleum virus 6 structure with a five-turns repeat (Tollin *et al.,* 1992).

The fact that closteroviruses possess the most flexible particles among elongated RNA viruses has been a key feature for their recognition as a distinct group (Brandes and Bercks, 1965). This apparent flexibility is reflected in a lower ratio of closterovirus RNA mass to modal particle length (2831–3230/nm) as compared with that for potex-, poty-, and carlaviruses (4038–4112/nm), or tobamoviruses (6666/nm) (Bar-Joseph and Hull, 1974; Bar-Joseph *et al.,* 1979). A loosely wound helix of BYV and CTV particles may account for their sensitivity to ribonuclease (Bar-Joseph and Hull, 1974).

It has been found that the BYV genome encodes a 24-kDa protein (24K) strikingly similar to the 22-kDa CP (22K), and that both these proteins have counterparts encoded in the CTV genome (Boyko *et al.,* 1992). It has been suggested that the genes for CP homologues arose by gene duplication that probably occurred in a common closterovirus ancestor; it is noteworthy that, despite significant divergence, the CP homologues of BYV and CTV have retained the profile of conserved amino acid residues that are believed to ensure the characteristic fold of the filamentous plant virus CPs (Boyko *et al.,* 1992, Dolja *et al.,* 1991). This discovery posed a question of the presence of a second structural protein in the virions. Initially, there were doubts as to whether this is theoretically possible, and researchers in this field (the author being no exception) have long been under the spell of the "single CP" paradigm of elongated plant virus structure (Boyko *et al.,* 1992; Dolja *et al.,* 1994). Then, quite unexpectedly, immunoelectron microscopy (IEM) revealed two serologically distinct segments in the BYV particles: a "tail" selectively labeled with antibodies to the 24-kDa protein (Fig. 1A), and the main segment labeled with antibodies to purified virus and to the recombinant 22-kDa protein (Fig. 1B; D.-E. Lesemann and author, unpublished observations). After decoration with antibodies to the 24-kDa protein, the BYV particles exhibited two peaks of modal lengths corresponding to tailed 1370-nm particles and to 1293-nm tailless particles; the lengths of the anti-24K serum-decorated tails peaked at 75 nm (Agranovsky *et al.,* 1995). The good balance between these lengths illustrates the propensity of the tails to break off and the relative stability of the 22K-coated part against further degradation. In line with this, purified BYV preparations contained a lower portion of the tailed particles as compared with crude

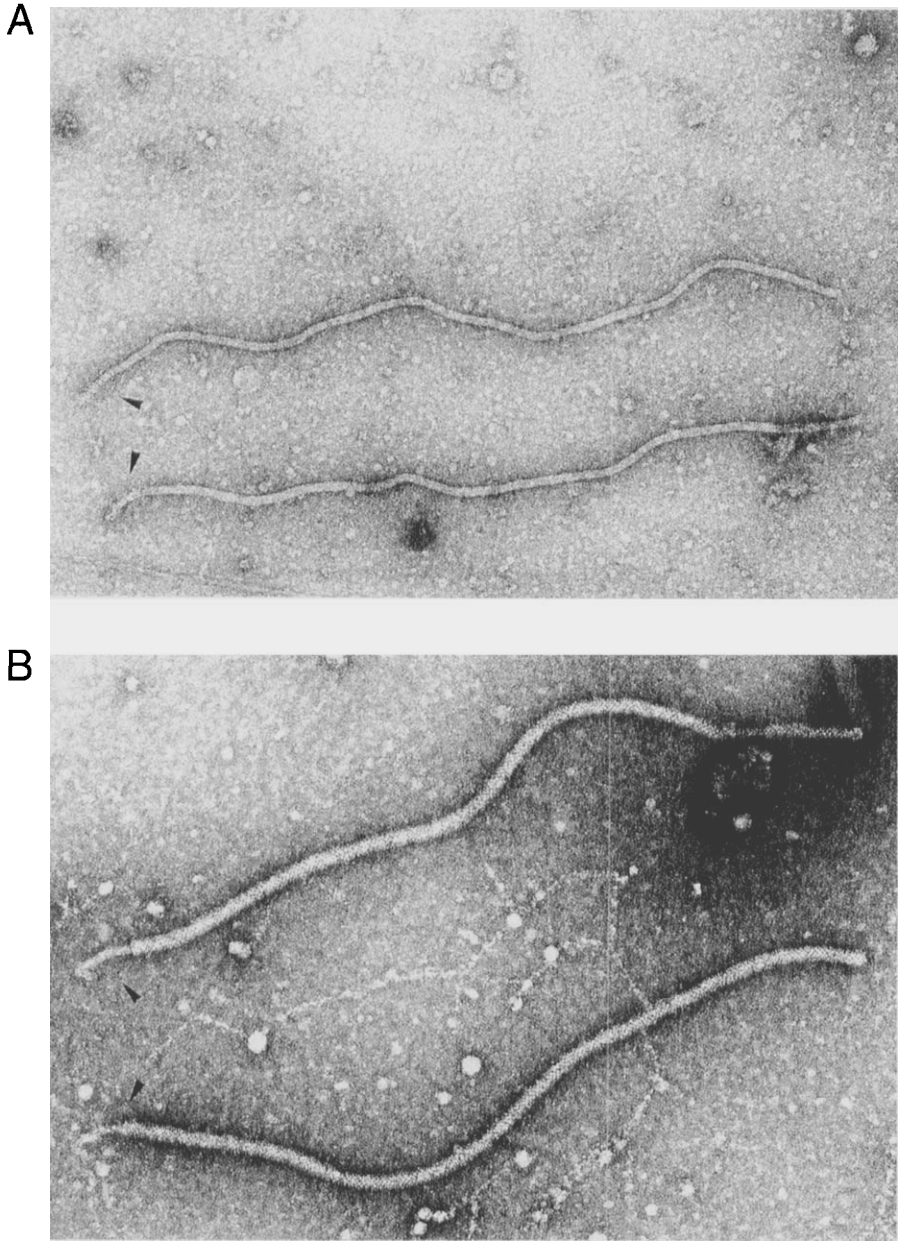

FIG. 1. Immunoelectron micrographs of BYV particles in crude sap of infected *Tetragonia expansa* plants (×95,000 magnification). (A) Decoration with mouse polyclonal antiserum against the N-terminal peptide of the BYV 24K protein. (B) Decoration with rabbit antiserum against purified BYV particles. Grids with the adsorbed virus particles were incubated with the antisera for 15 min (A) and 30 min (B), followed by negative staining with uranyl acetate. Arrows tag the distinct vinon tail. (Courtesy of D.-E. Lesemann.)

sap extracts, and the tails were no longer observed in the particles in the infected plant sap after overnight incubation at room temperature (D.-E. Lesemann, unpublished observations). The discrepancy between the BYV particle lengths determined for purified virus preparations and for leaf-dips (1250 versus 1370–1450 nm; reviewed in Bar-Joseph *et al.*, 1979) may also be at least partially attributed to preservation or loss of the tails.

Strange as it may seem, the morphological polarity of the BYV particles was first noticed much earlier. With the aid of electron microscopy of the methylamine tungstate-stained BYV particles, Hills and Gay (1976) observed an 83-nm terminal region with a helix pitch of 4.05 nm, which was clearly distinct from the main part of the 1390-nm virion having a pitch of 3.45 nm. These observations, which long remained unexplained, are consistent with the IEM data discussed above.

Thus, BYV particles, unlike those of other elongated plant viruses, possess a distinct tail built of multiple subunits of a minor CP, and hence have a "rattlesnake" rather than uniform structure (Agranovsky *et al.*, 1995). Comparison of the lengths of the 24K- and 22K-encapsidated segments (75 vs. 1293 nm) gives a rough estimate of one 24-kDa molecule per 17 molecules of 22K, which is consistent with the relative proportions of the respective subgenomic mRNAs in BYV-infected tissues (Dolja *et al.*, 1990). Moreover, the putative subgenomic promoters for the 22K and 24K mRNAs are similar (see Section VI,C), thus suggesting concerted expression of both structural proteins in viral infection. The involvement of 24K in virion formation is in line with previous computer predictions that the closterovirus CP duplicates have a spatial fold conserved in the monophyletic family of filamentous virus CPs (Boyko *et al.*, 1992; Dolja *et al.*, 1991). Specific decoration of the BYV particles with antibodies against the N-terminal peptide of 24K strongly indicates that the N terminus of the minor CP is exposed on the virion surface, as is the case with CPs of filamentous potex- and potyviruses (Agranovsky *et al.*, 1995; Koenig and Torrance, 1986; Shukla *et al.*, 1988).

It is possible that other closteroviruses have a similar virion structure. There is remarkable CP size heterogeneity in purified CTV preparations, albeit at least partly due to posttranslational modification of the (major) CP (Sekiya *et al.*, 1991; Lee *et al.*, 1988). Intriguingly, LIYV preparations purified in Cs_2SO_4–sucrose gradients contained minor amounts of an approximately 55-kDa protein along with the 28-kDa CP (Klaassen *et al.*, 1994). The putative CP duplicate encoded in the RNA-2 of this virus has a deduced molecular weight of 52 kDa (Fig. 2; Klaassen *et al.*, 1995).

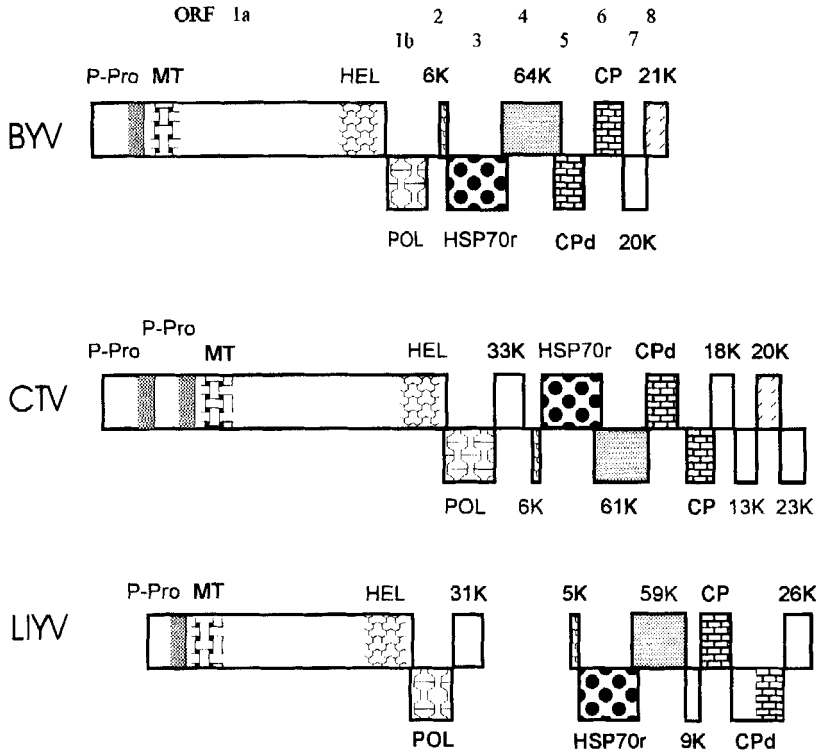

FIG. 2. Comparison of the genome maps of the closteroviruses BYV, CTV, and LIYV. ORFs are shown as boxes, with related domains indicated by the same fill-pattern. P-Pro, papain-like proteinase; MT, methyltransferase; HEL, helicase; POL, RNA-dependent RNA polymerase; HSP70r, HSP70-related protein; CP and CPd, capsid protein and its diverged duplicate.

How is the "rattlesnake" particle assembled? The principal mechanism might follow the classical scheme of TMV self-assembly, which starts at an internal origin of assembly (OAs) in the genomic RNA, and continues by adding CP disks or smaller aggregates in both the 5′ and 3′ directions (reviewed in Dobrov and Atabekov, 1989; Butler, 1984; Mathews, 1991; Lomonossoff and Wilson, 1985). We have assumed the existence of a nucleation region inherent in the BYV RNA that might discriminate between the capsid proteins or their disks, whereupon they proceed to encapsidate the RNA (Agranovsky *et al.,* 1995). In order to identify the BYV particle's end made up of 24K, we have run the following experiment: purified particles were sonicated and treated with antibodies to 24K, and the antibody–virion fragment complexes

were isolated on Protein A-Sepharose. The RNA extracted from the virion fragments thus selected was 5′-end labeled with [γ-^{32}P]ATP and used as a probe to develop Southern blots with cDNA clones representing the 5′-terminal, middle, and 3′-terminal regions of the BYV genome. Compared with the RNA probe prepared from nonfractionated virion fragments, the antibody-selected probe apparently hybridized more strongly to the 5′-terminal clones, thus suggesting that this was the 5′-terminal region of the BYV RNA associated with 24K, and that the putative nucleation signal might be mapped to this region (unpublished data). This is consistent with the fact that none of the 3′-coterminal subgenomic RNAs of BYV produced in infected plants are found in purified virus preparations (Dolja *et al.*, 1990). Closterovirus particle formation may be assisted by virus-encoded nonstructural protein(s); in some filamentous DNA phages carrying a few copies of minor CPs at their ends, assembly is chaperoned by phage-encoded proteins that are not part of the mature particles (reviewed in Russel, 1993). BYV encodes at least one nonstructural protein likely to be instrumental in protein–protein interactions, a 65-kDa homologue of the HSP70 cell heat-shock proteins (see Section V,A).

V. Genome Organization and Functions of Gene Products in Closteroviruses

A. Aphid-Transmissible Monopartite Closteroviruses

1. Beet Yellows Virus

The genome of the Ukrainian strain of BYV (BYV-U) consists of 15,480 nucleotides (nt), is 5′-capped and contains no 3′-poly(A) (Karasev *et al.*, 1989; Agranovsky *et al.*, 1991b, 1994a). Computer translation reveals nine ORFs in the sequence, flanked by 5′- and 3′-untranslated regions of 107 and 141 nt, respectively (Fig. 2). For the 3′ region, two potentially stable hairpins were predicted, and it has been speculated that these may serve as a recognition signal for viral replicase. BYV RNA cannot be aminoacylated or adenylylated *in vitro* and thus apparently has no 3′-tRNA-like structure (Agranovsky *et al.*, 1991a).

The 5′-proximal ORF1a codes for the 295-kDa product which encompasses the domains with methyltransferase and RNA helicase sequence motifs (MT and HEL; Fig. 2) that are conserved in the large subsets of positive-strand RNA viruses (Gorbalenya and Koonin, 1993; Rozanov *et al.*, 1992). The MT domain is believed to be involved in the capping of viral mRNAs. This activity has been experimentally demon-

strated for the alphavirus nsP1 protein (Mi and Stollar, 1991) and TMV 126-kDa protein (Dunigan and Zaitlin, 1990), and suggested for the closely related domains in the Sindbis-like supergroup virus replicases (Ahlquist *et al.*, 1985; Rozanov *et al.*, 1992). Likewise, the HEL domain, whose strand-separating activity was shown experimentally for the potyvirus CI protein (Lain *et al.*, 1990) and the pestivirus NS3 protein (Warrener and Collett, 1995), is implicated in unwinding of RNA duplexes on replication of many other virus groups, based on clear conservation of its sequence in the established and putative helicases (Gorbalenya *et al.*, 1988; Gorbalenya and Koonin, 1989, 1993; Hodgman, 1988). The ORF1b overlaps the last 40 triplets of ORF1a and codes for a product of approximately 53 kDa (Fig. 2) containing the domains of RNA-dependent RNA polymerase (POL) (Kamer and Argos, 1984; Koonin, 1991). Putative BYV replicase, which is likely to be expressed as an ORF1a/1b 348-kDa fusion protein (see below), should have a size about twice that of the other related viral replicases. This difference is due to two unique regions in the putative ORF1a/1b fusion product: a 600-residue N-terminal overhang containing a domain of cystein papain-like proteinase (P-Pro), and a 700-residue central insert harboring a 100-residue stretch that may be related to retrovirus aspartyl proteases (Agranovsky *et al.*, 1994a). The BYV P-Pro was found to be moderately similar to the C-terminal P-Pro domain in helper component proteases (HC-Pro) of potyviruses (Carrington *et al.*, 1989). HC-Pro proteases are multifunctional proteins required for potyvirus transmission by aphids (Atreya *et al.*, 1992; Pirone, 1991) and long-distance spread in plants (Cronin *et al.*, 1995). By analogy, similar function(s) may be proposed for the BYV leader proteinase.

The next downstream ORFs (2, 3, and 4) are arranged as an overlapping block and encode 6.4-, 65-, and 64-kDa products, respectively (Fig. 2). The BYV 6.4-kDa protein (6.4K) shows marginal similarity to the small hydrophobic proteins encoded in the "triple gene block" (TGB) of potex- and carlaviruses (Agranovsky *et al.*, 1991b; Morozov *et al.*, 1989). However, only a part of the residues conserved in the BYV 6.4K and the TGB-encoded proteins may be found in the approximately 6-kDa proteins encoded in CTV, beet yellow stunt virus (BYSV), and LIYV (Karasev *et al.*, 1994b; Klaassen *et al.*, 1995). Though their common origin is thus questionable, the small hydrophobic proteins of closteroviruses and potexviruses may be functionally equivalent. The TGB-encoded proteins of potexviruses bind to membranes *in vitro* (Morozov *et al.*, 1990) and mediate the cell-to-cell transport of the viral infection *in vivo* (Beck *et al.*, 1991). The BYV 6.4K synthesized in

rabbit reticulocyte lysates also showed affinity to cell membranes, and its involvement in the virus infection transport has been suggested (reviewed in Dolja *et al.,* 1994).

The 65-kDa protein (65K) is strikingly similar to the HSP70 family of cell heat-shock proteins (Agranovsky *et al.,* 1991a). HSP70s are ubiquitous molecular chaperones which assist proper folding, oligomerization, and transmembrane transport of other proteins (reviewed in Gething and Sambrook, 1992). Structurally, HSP70s consist of two parts, the N-terminal ATPase domain and the C-terminal peptide-binding domain (reviewed in Craig *et al.,* 1993). The BYV 65K protein contains an N-terminal domain whose sequence and tentative spatial fold are very similar to the HSP70 ATPase, and a unique C-terminal domain that cannot be folded into the $\beta_4\alpha$ (HLA-like) structure typical of the HSP70 peptide-binding domains (Agranovsky *et al.,* 1991a; Rippmann *et al.,* 1991; F. Rippmann, personal communication). Hence, the structure of the putative protein-binding domain of 65K suggests a function different from that of classical chaperones. Karasev *et al.* (1992) first reported that BYV 65K expressed in a cell-free transcription–translation system coprecipitates with purified bovine brain microtubules. The binding of 65K was abolished by pretreatment of microtubule preparations with subtilisin, thus suggesting its specificity. Very recently, bacterially expressed 65K and its fragments have been produced in our laboratory (Nikiphorova *et al.,* 1995). Using a polyclonal antiserum to the C-terminal 13-kDa fragment of 65K, the protein was detected in BYV-infected *Tetragonia expansa* plants (Agranovsky *et al.,* manuscript in preparation). *In vitro* assays showed that the purified BYV 65K, like the cell HSP70s, has magnesium-dependent ATPase activity associated with its N-terminal 40-kDa fragment. However, 65K, unlike its cell homologues, was found to be unable to bind to immobilized denatured protein, and its ATPase activity was not stimulated *in vitro* by sequence-nonspecific peptides (A. Agranovsky, S. Nikiphorova, O. Denisenko, and A. Folimonov, unpublished data). Although these data establish some biochemical characters of the BYV 65K pertinent to its function, the possible involvement of 65K in the cell-to-cell movement of the closterovirus infection (Agranovsky *et al.,* 1991a), which may involve specific interactions with the cell cytoskeleton and translocation machinery (Karasev *et al.,* 1992), awaits experimental support.

Internal segments in the BYV 64-kDa protein and in the equivalent CTV 61-kDa protein reportedly show similarity to a domain in the HSP90 heat-shock proteins (Koonin *et al.,* 1991; Pappu *et al.,* 1994).

However, the related approximately 60-kDa proteins of LIYV and SPSVV fail to display this similarity (Klaassen *et al.,* 1995; author, unpublished observation).

ORF5 and ORF6 code for the 24K and 22K capsid proteins of BYV, respectively (Agranovsky *et al.,* 1991b, 1995). The bacterially expressed BYV 24K and 22K share some common epitopes. Upon tissue fractionation, both proteins bulk in the soluble fraction of the BYV-infected cells, but are also found in the cell wall and membrane fractions (Agranovsky *et al.,* 1994b). The structure of the BYV virions built of two CPs is reminiscent of some other plant RNA viruses, thus implying functional analogy. First, one cannot but recall the rod-shaped furoviruses and spherical luteoviruses harboring a few copies of CP extended by readthrough of a leaky terminator codon in the CP gene (Bahner *et al.,* 1990; Cheng *et al.,* 1994; Filichkin *et al.,* 1994; Richards and Tamada, 1992). Notably, such an aberrant protein has recently been mapped to one end of furovirus particles (Haeberle *et al.,* 1994). The readthrough CP species are held to ensure the persistence of furoviruses and luteoviruses in their respective vectors: fungal zoospores and aphids. Likewise, the semipersistent mode of BYV transmission may be due to the ability of the assembled 24K to cling tightly to cell membranes lining the aphid's alimentary tract. Another (and not necessarily alternative) possibility may be that the 24K tail directs the closterovirus particle to a host (phloem) cell receptor. Conceivably, the fact that p24 is involved in formation of mature virions does not discredit the earlier suggestion that it might participate in the formation of nonvirion ribonucleoproteins adapted for the cell-to-cell transport of genomic RNA (Boyko *et al.,* 1992; Dolja *et al.,* 1994).

ORFs 7 and 8 encode 20- and 21-kDa products, respectively. The latter is related to a 20-kDa protein encoded in the CTV genome (Pappu *et al.,* 1994). Apart from this, these products have shown no significant similarities to any proteins in the current database. Recently we produced a polyclonal antiserum against the BYV 21-kDa protein purified from bacteria; using this antiserum to develop Western blots, the 21-kDa protein was detected in soluble and membrane fractions of BYV-infected plants (R. Zinovkin and author, unpublished data).

The German and British strains of BYV (BYV-G and BYV-B) have been partially sequenced (Agranovsky *et al.,* 1994a; Brunstedt *et al.,* 1991), allowing their comparison with BYV-U. BYV-U and BYV-G showed 88.5% identity of the nucleotide sequences and the same organization of ORFs 2 to 8 within the 3′-terminal 6-kb region. The majority of nucleotide substitutions in the BYV-G sequence are in the third

positions of codons; even when the substitutions change the coding, only about half of the resulting amino acid changes are nonconservative. Nevertheless, the data compiled in Table II indicate some differences in the extent of conservation of individual protein sequences in the two BYV strains. Proteins 65K, 64K, 24K, and 22K are the best conserved among the strains, whereas the low-molecular-weight proteins (6.4K, 20K, and 21K) are apparently more variable (Table II). There is also a remarkable nucleotide sequence conservation of the intergenic and 3′-untranslated regions among the two strains, suggesting the functional importance of these regions (Table II). The partial sequence of BYV-B shows the same disposition of ORFs 4 to 7 (Brunstedt *et al.,* 1991). The West European strains are apparently closer to each other than to the Ukrainian strain; in particular, the 22-kDa capsid protein sequences of the BYV-G and BYV-B are identical (Table II).

TABLE II

COMPARISON OF THE NUCLEOTIDE AND AMINO ACID SEQUENCES AMONG THE THREE BYV STRAINS

Gene product	ORF2 64K	ORF3 65K	ORF4 64K	ORF5 24K	ORF6 22K	ORF7 20K	ORF8 21K
Strain							
U/G[a]	11[b]	4	8	2	5	18	12
U/B	—	—	7	2	5	—	—
G/B	—	—	2	2	0	—	—
Nontranslated regions			I		II		3′ end
Strain							
U/G			14		14		4
U/B			—		13		—
G/B			—		1		—

[a] U, G, and B stand for the Ukrainian, German, and British strains of BYV, respectively. The available nucleotide sequences of the BYV-G and BYV-B align with nt 9375–15353 and 11684–14407, respectively, in the complete BYV-U sequence.

[b] Percent of amino acid (for polypeptide products) or nucleotide (for noncoding regions) substitutions revealed on pairwise comparisons of strains. Dashes indicate positions where no sequence for the British strain was available. Nontranslated region 1 is between ORFs 1b and 2, and nontranslated region II is between ORFs 5 and 6.

2. *Citrus Tristeza Virus*

Among plant RNA viruses, CTV has the largest undivided genome (19,296 nt for the Florida T36 isolate), exceeding that of BYV by about 4 kb (Karasev *et al.*, 1994b, 1995; Pappu *et al.*, 1994). The overall genome structure of CTV is similar to that of BYV, comprising the P-Pro, MT, HEL, and POL domains; the small hydrophobic protein; the HSP70 homologue; the 61-kDa protein related to BYV 64K; the 27-kDa CP homologue (27K); the 25-kDa CP; and the 20-kDa protein homologous to the BYV 21-kDa protein (Fig. 2). On the other hand, the CTV genome encodes some proteins or polyprotein domains that are not conserved in BYV. Interestingly, the P-Pro domain in the ORF1a product of CTV is duplicated (Fig. 2; Karasev *et al.*, 1995). Pairwise comparisons of the putative CTV leader proteins of predicted molecular weights 54 and 55 kDa and the BYV 66-kDa protein revealed no sequence similarity among the three proteins apart from the C-terminal 150-residue part encompassing the P-Pro domain. Among the products encoded by the 3′-proximal genes of CTV, the 20-kDa protein (20K) is related to the BYV 21-kDa protein (21K), whereas the 33-, 13-, 18-, and 23-kDa proteins have no homologues in other sequenced closterovirus genomes (Karasev *et al.*, 1995; Pappu *et al.*, 1994). The 23-kDa protein contains a sequence motif enriched in cysteine and basic residues, which is conserved in putative nucleic acid binding proteins encoded in the 3′-proximal genes in carlaviruses and allied viruses. Therefore, this putative protein has been implicated in RNA binding and regulation of CTV gene expression (Dolja *et al.*, 1994).

Severely pathogenic CTV isolates share a common epitope on their particles not found on the particles of mild isolates, thus suggesting the CTV pathogenicity may have some of its determinants associated with the 25-kDa CP (Pappu *et al.*, 1993). The CTV 27K and 20K proteins were detected in infected plants with polyclonal antibodies against the recombinant proteins (Febres *et al.*, 1994; Pappu *et al.*, 1994). The bulk of 27K was found in the cell wall fraction of infected citrus leaves, although the protein was also detectable in the soluble and membrane fractions (Febres *et al.*, 1994). Thus, the 27K association with subcellular fractions differs somewhat from that reported for the homologous BYV 24K which, like BYV 22K, is predominantly found in the soluble fraction (Agranovsky *et al.*, 1994b). Recent experiments with yeast two-hybrid system have shown that the CTV 20K (a homologue of the BYV 21K) is capable of homologous interactions, thus suggesting that this protein might function as a di- or multimer (S. Gowda, personal communication).

Sequencing of a 3′-terminal 2.5-kb portion of another CTV isolate, Israeli VT or "seedling yellows" isolate, revealed four ORFs encoding 18-, 13-, 20-, and 23.5-kDa proteins that showed close relatedness to the respective products of CTV-T36 (Mawassi *et al.*, 1995a).

3. Carnation Necrotic Fleck and Beet Yellow Stunt Viruses

Comparisons of partial sequences of BYSV and carnation necrotic fleck virus (CNFV) indicate their relatedness to CTV and BYV (Dolja *et al.*, 1994; Karasev *et al.*, 1994a,b; Klaassen *et al.*, 1995). Both the BYSV and CNFV genomes contain a conserved array of ORFs coding for POL, a small hydrophobic protein, and an HSP70 homologue. In addition, the BYSV genome bears an approximately 30-kDa protein gene inserted between the POL and small hydrophobic protein genes. This configuration is similar to that in the respective part of the CTV genome (Fig. 2). The fact that BYV and CNFV induce very similar patterns of dsRNAs in the infected plant cells (Dodds and Bar-Joseph, 1983) is indicative of the overall similarity of their gene layouts.

B. Whitefly-Transmissible Mono- and Bipartite Closteroviruses

The genome of lettuce infectious yellows virus (LIYV) is divided among RNA-1 and RNA-2 components of 8.1 and 7.2 kb, respectively (Klaassen *et al.*, 1995). Interestingly, LIYV RNA-1 and RNA-2 show no similarity between their respective terminal untranslated regions (which may be expected to contain putative recognition signals for the replicase), with the exception of the 5′-terminal pentanucleotide, which is identical in both genomic components (Klaassen *et al.*, 1995). LIYV RNA-1 encompasses the overlapping ORFs 1a and 1b coding for the putative P-Pro and the replicative domains, and the 3′-terminal ORF for a 31-kDa protein (Fig. 2). Very recently, a full-length cDNA copy of the LIYV genome was produced, and it was found that the RNA-1 T7 transcript is necessary and sufficient to support the replication in protoplasts (B. W. Falk, personal communication). This is the first experimental evidence for the assignment of closterovirus replicative functions to the domains conserved in ORF 1a/1b. LIYV RNA-2 contains genes for the small hydrophobic protein, the HSP70 homologue, the 59-kDa protein distantly related to the BYV 64-kDa and CTV 61-kDa products, the 9-kDa protein, the 28-kDa CP, the 52-kDa protein whose C-terminal domain is homologous to the CP, and the 26-kDa protein (Fig. 2; Klaassen *et al.*, 1995). SPSVV RNA-1 and RNA-2 have sizes comparable to those of their counterparts in the genome of LIYV and show a similar organization of ORFs, with the biggest difference being

that SPSVV RNA-2 encodes a putative CP duplicate of 79-kDa (U. Hoyer, E. Maiss, W. Jelkmann, and J. Vetten, unpublished data).

The monopartite genome of another whitefly-transmissible closterovirus, cucumber chlorotic spot virus (CCSV), has a size of approximately 15.5 kb (Woudt *et al.*, 1993a,b). The sequence of its coding part shows 5′-terminal overlapping ORFs encoding the domains of the leader P-Pro, MT, HEL, and POL; 3′-proximal ORFs code for the small hydrophobic protein, the HSP70 homologue, the approximately 60-kDa protein, the 9-kDa protein, the 28-kDa CP, the 74-kDa protein containing the C-terminal domain homologous to the CP, and the 23-kDa protein (L. P. Woudt, personal communication). Thus, the undivided genome of CCSV shows an overall arrangement of genes unexpectedly similar to that of the bipartite genomes of LIYV and SPSVV. In line with this, comparisons of the encoded proteins showed close relatedness among CCSV, LIYV, and SPSVV, suggesting that the three whitefly-transmissible closteroviruses constitute a distinct evolutionary lineage (Fig. 3). The LIYV, SPSVV, and CCSV ORF 1a and 1b products (including the N-terminal leader proteins) can be confidently aligned with high statistical scores over almost the entire protein length, whereas their similarity to the equivalent products of BYV and CTV is essentially confined to the core replicative domains. The same is true for the encoded HSP70 homologues, the approximately 60-kDa proteins, and the 28K-kDa CPs. The LIYV 52-kDa, SPSVV 79-kDa, and CCSV 74-kDa capsid protein duplicates show closest relatedness within the approximately 200-residue C-terminal segments, including the CP-like core domains. The gene for a putative 9-kDa protein located upstream of the CP gene is unique for the whitefly-borne closteroviruses; comparison of the 9-kDa sequences encoded in LIYV, SPSVV, and CCSV showed moderate conservation. The presence of a gene for an approximately 30-kDa product located downstream of the POL gene is common for CTV, BYSV, and the bipartite closteroviruses, although the relatedness of the encoded products is not apparent (Dolja *et al.*, 1994; Karasev *et al.*, 1994b; Klaassen *et al.*, 1995). The CCSV genome does not encode a product of similar size and location.

C. Other Closteroviruses

The 3′-terminal 8.3-kb sequence of the mealybug-transmissible LCV has been recently determined (R. Keim-Konrad and W. Jelkmann, manuscript in preparation). In the 5′ to 3′ direction, the sequence encompasses the conserved ORFs for the HSP70 homologue, the ~60K protein, the 46-kDa (putative) CP, the 76-kDa CP duplicate, and

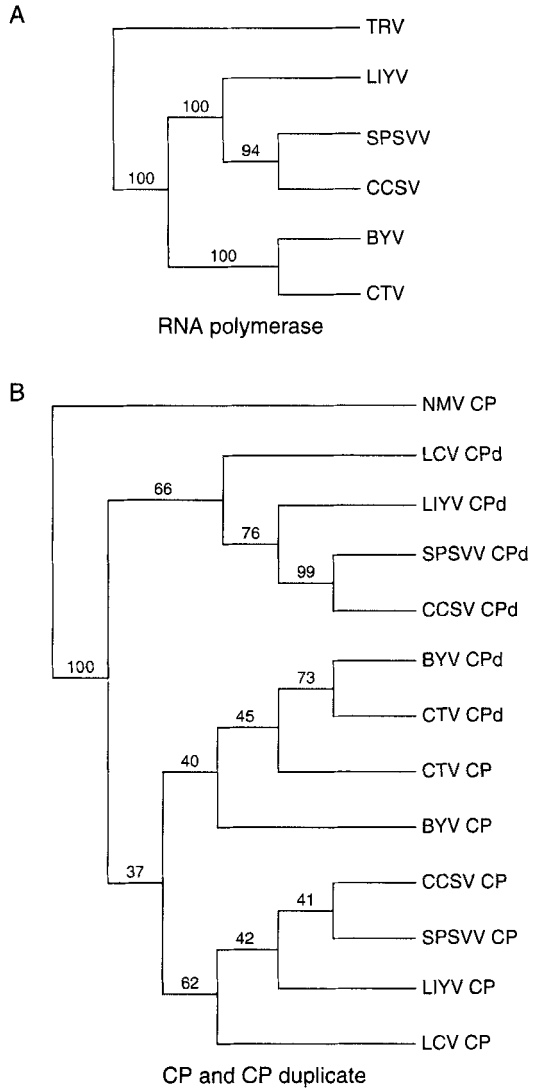

FIG. 3. Tentative phylogenetic reconstructions derived from aligned sequences of RNA-dependent RNA polymerase (A) and capsid protein core (B) domains of closteroviruses. Multiple alignments were produced by the program OPTAL (Gorbalenya *et al.*, 1989). Trees were constructed using the program PROTPARS of the PHYLIP package (Felsenstein, 1989) from the alignments of 294 and 98 amino acid residues of polymerases and CPs, respectively, excluding the positions containing gaps. The number above each node shows the percentage of bootstrap replicates in which a given node was recovered. Branch lengths are arbitrary. The related sequences of tobacco rattle tobravirus (TRV) polymerase and narcissus mosaic potexvirus (NMV) CP were used as the outgroups in the respective trees. Protein sequences were extracted from database, except for CCSV, SPSVV, and LCV (personal communications from L. P. Woudt, U. Hoyer, and W. Jelkmann, respectively).

unique ORFs for 21- and 27-kDa proteins (W. Jelkmann, personal communication). The genome of GLRaV-3 reportedly contains the HEL and POL domains, a gene for the HSP70 homologue, and the 43-kDa CP gene (Ling *et al.,* 1994), but their arrangement has not been yet described. Future sequencing and comparison efforts are expected to elucidate relationships among mealybug-transmissible closteroviruses.

D. General Outlook of Genome Structure and Gene Functions

In closterovirus genome organization, variation of a common theme is evident. The invariant elements include the P-Pro, MT, HEL, and POL domains; the small hydrophobic protein; the HSP70 homologue, the ~60K protein; and the CP and its duplicate. Interestingly, the 65K, 64K and CP genes that are conserved at the species level are apparently least divergent in the BYV strains. It is plausible that these genomic elements provide for the characteristic biological patterns common to all the members of this family. Wherever differences in the conserved genes occur, they intriguingly parallel modifications of the biological features. For example, the genes for the (major) CP and its duplicate are much more divergent and are transposed in the genomes of whitefly-transmissible CCSV, LIYV, and SPSVV as compared to aphid-transmissible BYV and CTV. In the genome of a mealybug-transmissible LCV, the CP and CP-duplicate ORFs are arranged like those in the whitefly-transmissible species; however, their amino acid sequences are divergent from those of the other family members (Fig. 3), and the LCV CP is notably large. Could the disparity among the closterovirus CPs and CP duplicates be due to their involvement in and specific adaptation to transmission by different types of vector?

The obvious similarities in gene arrangement and the encoded protein sequences between the whitefly-borne monopartite and bipartite closteroviruses clearly indicate that the lineages represented by CCSV, LIYV, and SPSVV on the one hand, and by BYV and CTV on the other, had diverged at an evolutionary stage preceding the splitting of the closterovirus genomes. This has important implications in the taxonomy of *Closteroviridae*; clearly, the second genus (Table I) should include not only the bipartite closteroviruses but also CCSV (and perhaps other whitefly-transmitted closteroviruses). Broadly, sequencing of more closterovirus genomes is expected to shed some light on whether the adaptation to vectors has been a key factor of their molecular evolution.

In closterovirus genomes, the 3′-proximal ORFs vary in number and, as a rule, encode nonconserved protein sequences. The 3′-most ORFs

have perhaps diverged most rapidly among the BYV strains (Table II). This might indicate that their products are involved in functions connected with fast environmental response, such as modulation of symptom expression or adaptation of the virus to changes in the host or vector populations. A number of BYV and CTV isolates have been reported, differing in the severity of symptoms they cause in host plants (reviewed in Moseley and Hull, 1990; Pappu *et al.*, 1993; Rogov *et al.*, 1993).

Closterovirus genomes appear to have modular organization (reviewed in Dolja *et al.*, 1994). In the beginning, the existence of three modules in the BYV genome was envisaged, represented by overlapping gene blocks separated by two uridine-rich spacers (Agranovsky *et al.*, 1991b). It has been suggested that these modules have evolved as distinct entities and that they encode proteins expressed very early (replicase), early (putative transport proteins), and late in the infection (CP and two proteins of unknown function). Revisions of this scheme have been proposed (Dolja *et al.*, 1994; Karasev *et al.*, 1995); they agree in placing the genes for the CP and its duplicate into the 3′-terminal module, which seems reasonable. Further, it may be speculated that some closterovirus genes (the approximately 30-kDa ORF and the unique 3′-terminal ORFS) evolved independently of the conserved modules.

VI. Expression Strategy

A. Papain-like Proteinase

Computer-assisted predictions and *in vitro* experiments have demonstrated that the 295-kDa product encoded in the 5′-most gene of BYV is in fact a polyprotein (Agranovsky *et al.*, 1994a). The P-Pro domain located in its N-terminal portion mediates autoproteolysis at the Gly–Gly bond to release a 588-residue (66-kDa) leader protein and a C-terminal 229-kDa protein with MT and HEL domains. The catalytic Cys and His residues in the P-Pro active center and its cleavage site (inferred from alignment with potyviral P-Pro domains) have been confirmed experimentally using point mutagenesis and *in vitro* translation. It has been found that deletion of 245 residues from the N terminus of the leader protein does not impair but rather stimulates the cleavage, and that the His and Cys residues, which are not conserved in the related thiol proteinases, have different effects on autoproteolysis of the BYV polyprotein (Agranovsky *et al.*, 1994a). This is consistent with the results obtained for a related P-Pro of the chestnut blight hypovirulence-associated dsRNA virus (Choi *et al.*, 1991).

In the CTV ORF1a product, the P-Pro domain is duplicated (Karasev *et al.,* 1995). Based on alignment of the CTV and BYV P-Pro sequences, the cleavage sites in the CTV polyprotein have been predicted at the Gly–Gly doublets at positions 484–485 and 976–977 from the N terminus. Hence, processing of the 349-kDa ORF1a product of CTV would yield two leader proteins of 54 and 55 kDa, and a 240-kDa (C-terminal) protein with MT and HEL domains (Karasev *et al.,* 1995). Putative leader P-Pro domains may also be revealed on computer analysis of the ORF1a products of LIYV (Klaassen *et al.,* 1995), SPSVV (U. Hoyer, personal communication), and CCSV (L. P. Woudt, personal communication). The tentative P-Pro cleavage sites in the ORF1a products of LIYV, SPSVV, and CCSV deviate from the consensus drawn for the BYV, CTV, and potyvirus proteinases, being VG/A, LG/V, and VG/V, respectively. If autocatalysis indeed occurs at these sites, the respective leader proteins should have 412, 496, and 402 residues, respectively.

B. Ribosomal Frameshifting

In the genomes of BYV (Agranovsky *et al.,* 1994a), CTV (Karasev *et al.,* 1995), CCSV (ten Dam, 1995), and bipartite closteroviruses (Klaassen *et al.,* 1995; U. Hoyer, E. Maiss, W. Jelkmann, and J. Vetten, unpublished data), the HEL and POL domains are split between the products of overlapping 5′-proximal ORFs found in 0/+1 configuration, thus indicating that the polymerase may be expressed via +1 ribosomal frameshifting. Although many viral RNA polymerases are expressed as frameshift fusions resulting from translation of overlapping genes, the 0/+1 configuration of the closterovirus replication-associated ORFs is quite unusual [to my knowledge, the only other example is the dsRNA virus of *Leishmania* (Stuart *et al.,* 1992)]. In all other cases, including retroviruses (Jacks and Varmus, 1985; Jacks *et al.,* 1988), dsRNA-containing viruses (Dinman *et al.,* 1991), and the diverse groups of positive-strand RNA viruses of animals and plants (Brierly *et al.,* 1987; Godeny *et al.,* 1993; Jiang *et al.,* 1993; Mäkinen *et al.,* 1995; Miller *et al.,* 1988; Xiong and Lommel, 1989), an upstream ORF and the downstream (POL) ORF are found in 0/–1 configuration. Hence, the tentative frameshifting mechanism in closteroviruses deserves a special comment.

The canonical mechanism of leftward (or –1) frameshifting postulates a one-step-back movement ("simultaneous slippage") of two tRNAs bound to a "shifty" mRNA sequence X XXY YYZ, to decode it as XXX YYY (Jacks *et al.,* 1988). The reading-frame switching is stimulated by a pseudoknotted secondary structure (Brault and Miller, 1992; Prüfer *et al.,* 1992; ten Dam *et al.,* 1990); this effect is probably con-

nected with the ability of such a structure to impede the progress of ribosomes along the template (Tu *et al.,* 1992). Rightward (or +1) frameshifting has been described for the yeast *Ty* retrotransposons and the *E. coli* release factor gene (Clare *et al.,* 1988; Craigen *et al.,* 1985; reviewed in Farabaugh, 1993). In retrotransposons, the frameshifting is enhanced by a rare "hungry" codon adjacent to the shifty codon (Farabaugh *et al.,* 1993), whereas in the bacterial gene stimulation is provided by the in-frame stop codon and a downstream Shine-Daigarno-like sequence transiently interacting with ribosomal 16S RNA (Weiss *et al.,* 1988). In short, as follows from comparisons of different frameshifting mechanisms, the reading-frame switching requires some signal(s) to slow down the translating ribosome, thus increasing the chances of the out-of-frame triplet recognition.

The BYV ORF1a ends in a GGGUUUA sequence resembling the "shifty" heptamers of the retroviral type. This resemblance, which we could not but mention in an earlier work (Agranovsky *et al.,* 1994a), is probably fortuitous, as such a heptamer is not conserved in the other closterovirus genomes (Fig. 4; Karasev *et al.,* 1995; Klaassen *et al.,* 1995; ten Dam, 1995). Notwithstanding, we did not suggest the "slippery" consensus to provide for the +1 frameshifting in the BYV svstem; rather, our explanation was based on the "U33 grapple" pairing model (Weiss, 1984). Specifically, offset pairing was postulated between U-7998 in the ORF1a UAG stop codon and the nucleotide located leftward to the anticodon of $tRNA^{Val}$ to mediate transition of a subset of translating ribosomes into ORF1b (Agranovsky *et al.,* 1994a). In accord with this, the (G/C)UU U** consensus (where ** designates the last two bases in stop codons) is seen at the 3′ termini of ORF1a in BYV, LIYV, SPSVV (Fig. 4), and CCSV (ten Dam, 1995). Further, it may be speculated that putative secondary structure elements at the BYV ORF1a stop codon (Agranovsky *et al.,* 1994a) serve to stall the ribosome, thus promoting the frameshifting. At least partially, this RNA fold is conserved in the respective genome regions of CTV and CCSV (Fig. 5; ten Dam *et al.,* 1995), but not in the LIYV genome (Klaassen *et al.,* 1995).

Alternative frameshifting models have been proposed for CTV and BYV (Karasev *et al.,* 1995) and for LIYV (Klaassen *et al.,* 1995). Superposition of the nucleotide and protein sequences in the CTV and BYV HEL/POL gene overlaps reveals a remarkable amino acid conservation profile, suggesting that the frameshifting in the CTV gene occurs after the GUU valine codon, which is not the penultimate triplet there (Fig. 4; Karasev *et al.,* 1995). By analogy with the yeast retrotransposon system (Farabaugh *et al.,* 1993), the putative +1 frameshifting signal is postulated to be simple, not to include any secondary

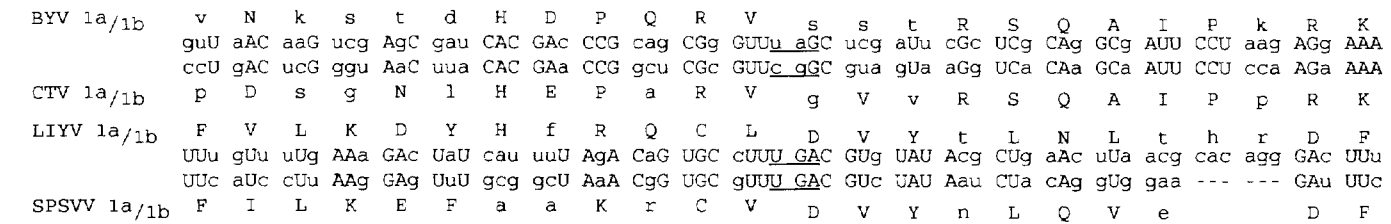

FIG. 4. Alignment of the nucleotide and amino acid sequences in the region of the putative ribosomal frameshift in the genomes of BYV, CTV, LIYV, and SPSVV. Residues conserved in two or more sequences are capitalized. Stop codons in the BYV, LIYV and SPSVV ORF1a and the respective Arg codon in the CTV ORF 1a are underlined. The ORF1a of CCSV ends in the CGG CGA GUUUGA sequence (L. P. Woudt, personal comrnunication) that is similar to those in LIYV and SPSVV.

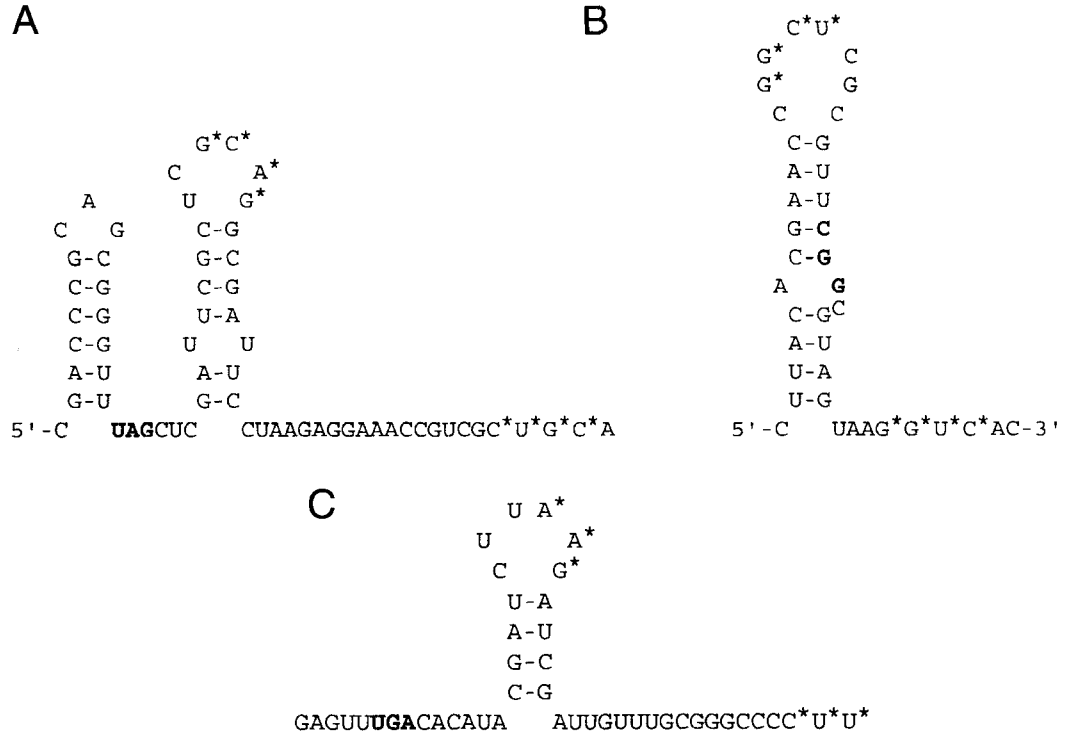

FIG. 5. Proposed structural fold for the regions of the putative ribosomal frameshift in the genomes of BYV (A), CTV (B), and CCSV (C). The UAG stop codon in the BYV ORF1a and the respective CGG arginine codon in the genome of CTV are shown in bold. Nucleotides potentially involved in pseudoknot formation are marked by asterisks. CCSV fold has been proposed by ten Dam (1995).

structure, and to use codons that may cause ribosome pausing, namely the rare CGG (arginine) codon in the CTV ORF1a or UAG stop codon in the BYV RNA (Dolja *et al.*, 1994; Karasev *et al.*, 1995). For LIYV, Klaassen *et al.* (1995) proposed that the +1 frameshift occurs by slippage of $tRNA^{Lys}$ on the AAAG string located eight triplets upstream of the stop codon in the ORF1a.

Experimental evidence has been obtained for ribosomal frameshifting on expression of closterovirus ORFs 1a and 1b, and we may not have to wait long for elucidation of the frameshift mechanism. An expression cDNA clone was produced, which contained, in a heterologous context, a 113-nt CCSV-specific fragment encompassing the potential GUU UGA "shifty stop" sequence and a tentative pseudoknot downstream (ten Dam, 1995). Translation of the SP6 transcript of this clone in wheat germ and rabbit reticulocyte-cell-free systems yielded ^{35}S-methionine-labeled products consistent with the expression of CCSV ORF1b via ribosomal frameshifting with an efficiency of about 2%. It is worth mentioning that the minimal "shifty stop" sequence of CCSV was found to be incapable of frameshifting *in vitro*, suggesting that a more elaborate signal must be involved (ten Dam, 1995). Likewise, we have produced an expression clone containing the BYV-specific insert encompassing the ORF1a/1b overlap. Translation of the T7 transcript of this clone in rabbit reticuloctye lysate resulted in ribosomal frameshifting with an efficiency of less than 1% (Agranovsky, Zelenina, and Morozov, unpublished data).

C. Subgenomization of 3′-Proximal Genes

Closterovirus genes located 3′-ward of the POL gene are likely to be expressed via formation of 3′-coterminal subgenomic (sg) RNA species. Plants infected with CTV, BYV, CNFV, SPSVV, LIYV, BPYV, LCV, and other clostero-like viruses contain a variety of dsRNA species, of which some may correspond to the subgenomic size messengers (Coffin and Coutts, 1992; Dodds and Bar-Joseph, 1983; Gunasinghe and German, 1989; Hu *et al.*, 1990; Larsen *et al.*, 1991; Namba *et al.*, 1991; Winter *et al.*, 1992; K. Eastwell, personal communication).

In BYV-infected plants, six species of double- and single-stranded RNAs have been identified by Northern blot hybridization, corresponding to the full-sized genomic RNA and to sgRNAs for the 65, 64, 24, 22, and 21K ORFs. The identity of the BYV sgRNAs for the 24, 22, and 21K ORFs is supported by *in vitro* translation of the respective dsRNA species denatured with methyl mercuric hydroxide, which yielded proteins compatible in size with those deduced for the ORF 5, 6, and 8

products; the product of the most abundant 1.6-kb dsRNA (corresponding to the sgRNA for 22-kDa CP) was immunoprecipitable with an antiserum to BYV particles (Dolja *et al.*, 1990). None of the BYV sgRNAs was found to be encapsidated (Dolja *et al.*, 1990).

The dsRNA patterns produced in CTV-infected plants vary greatly from strain to strain (Dodds *et al.*, 1987; Guerri *et al.*, 1991; Moreno *et al.*, 1990). For at least one particular CTV isolate, T36, this pattern was found to be stable when the virus had been propagated in different citrus hosts (Hilf *et al.*, 1995). Comprehensive Northern blot analysis of single- and double-stranded RNAs from CTV-T36-infected plants demonstrated the presence of nine 3′-coterminal sgRNA species representing the 33, 65, 61, 27, 25, 18, 20, and 23K ORFs (Hilf *et al.*, 1995). The 3.2-kb (CP) sgRNA of CTV-T36 (Hilf *et al.*, 1995), as well as the 3.2-, 1.6-, and 0.9-kb sgRNAs coding, respectively, for the CP, 20-kDa, and 23-kDa proteins of CTV-VT (Mawassi *et al.*, 1995a), were found to be encapsidated. In addition, the CTV-VT encapsidates a 2.4-kb RNA that possesses properties of a defective RNA; as revealed by sequencing, this species is composed of 1.1-kb and 1.3-kb regions derived from the 5′ and 3′ termini of the CTV genome (Mawassi *et al.*, 1995b). Conceivably, the presence of this defective RNA in virions implies that the putative origin of assembly is in the outskirts of the CTV genome (Mawassi *et al.*, 1995b).

The 5′ termini of the BYV sgRNAs for the major and minor CPs (22 and 24K) were mapped by primer extension to the adenosine residues found 52 and 105 nt upstream of the respective initiating codons (Agranovsky *et al.*, 1994b). The sequence at the starts of both sgRNAs of BYV is conserved (CCAUUUYA; Y for pyrimidine) and may thus represent a core element of the subgenomic promoter. Interestingly, this element resembles the sequences at the 5′ ends of the CP sgRNAs of tobamoviruses and *Bromoviridae* family members. Bearing in mind that the BYV repilcase is most closely related to those of the tobamolike viruses, it is tempting to speculate on parallel conservation, in the process of evolution, of template-binding domains in viral replicases and the signals they recognize in viral RNAs (Agranovsky *et al.*, 1994b). Recently, we mapped the 5′ end of the BYV sgRNA for the 64-kDa protein at the adenosine residue located 141 nt upstream of the ORF4 initiating codon (M. Vitushkina and author, unpublished data). The sequence at the respective start site, ACAUAAUU, significantly deviates from the consensus derived for the 22 and 24K sgRNAs. This, together with the fact that no sequence elements conforming to the CCAUUUYA consensus can be seen in the BYV genome sequence upstream of the AUG codons in ORFs 2, 3, 4, 7, and 8, suggests that the

BYV replicase may interact with different types of subgenomic promoters, thus providing for transcriptional regulation of closterovirus gene expression.

D. Peculiarities of Closterovirus Genome Expression

For genome expression, BYV and possibly other closteroviruses combine autoproteolysis by a papain-like proteinase, ribosomal frameshifting, and sgRNA formation, thus resembling the animal viruses belonging to the corona-like superfamily rather than any other known plant virus group (Agranovsky *et al.*, 1994a). The situation with BYV and CTV, whose genomes contain single and double P-Pro domains, respectively, further parallels that in coronaviruses and arteriviruses, some of which show similar P-Pro duplication (Godeny *et al.*, 1993; Karasev *et al.*, 1995; Lee *et al.*, 1991: Snijder and Horzinek, 1993). As closteroviruses and corona-like viruses represent evolutionarily disparate lineages (Koonin, 1991; Koonin and Dolja, 1993), it seems plausible that similar expression strategies in these groups have evolved independently to confer an advantage in expression of large RNA genomes.

Expression of the 5′-proximal genes in closterovirus genomes should produce some proteins in unequal amounts. Thus, translation of the BYV genomic RNA should yield the major 295-kDa protein and a fusion 348-kDa protein processed into the 66-, 229-, and 282-kDa proteins. In CTV, translation should result in 349- and 401-kDa polyproteins further processed into the 53- and 54-kDa cleaved leaders, and 240- and 290-kDa proteins (Karasev *et al.*, 1995). The synthesis of closterovirus ORF1a/1b fusion proteins (containing the complete array of replication-associated domains) is perhaps down-regulated (see Section VI,B), as is the case with other virus systems employing translational frameshifting (Brault and Miller, 1992, and references therein).

In many positive-strand RNA virus genomes, one can discern a trend to regulate the expression of POL and other replication-associated domains (1) by using a leaky nonsense codon or a frameshift signal to isolate the sequence coding for POL from the upstream coding sequence, and (2) by expressing the POL, MT, and HEL (or MT+HEL) domains as distinct products resulting either from polyprotein processing or from translation of individual genomic RNAs. Splitting of the viral replicase into distinct components, whose expression may be regulated separately, is likely to provide the required flexibility in performing different enzymatic functions in RNA replication, namely, unwinding of duplexes, asymmetric synthesis of (+) and (–) strands, synthesis of subgenomic RNAs, and RNA capping (Agranovsky *et al.*,

1994a). Unlike closteroviruses, all other plant viruses that utilize frameshift for POL expression have small genomes and encode neither MT nor HEL (Koonin and Dolja, 1993), whereas in corona-like virus genomes both the POL and HEL domains (in this order) are located 3′-ward of the frameshift site (Snijder and Horzinek, 1993). Thus, closteroviruses are the only viruses known so far in which the frameshift occurs between the sequences coding for HEL and POL.

VII. Evolution of Closterovirus Genomes

A. RNA Genome and RNA Replicase: Coordinated Evolution?

Comparisons of the MT, HEL, and POL sequences reveal close similarity of closteroviruses to tobamo-, tobra-, furo-, hordei-, idaeo-, bromo-, and ilarviruses, which comprise a compact "tobamo" lineage (Koonin and Dolja, 1993) within the Sindbis-like supergroup of positive-strand RNA viruses (Goldbach *et al.,* 1991). Apart from the conserved replicative core that has been vertically inherited from an ancestor shared with tobamo-like viruses, closterovirus genomes show elements of most probably horizontal acquisition (Fig. 6). This concerns the 65-kDa protein evidently homologous to the HSP70 family of cell chaperones, the CP and its duplicate, and (less likely) the leader P-Pro related to the potyvirus HC-Pro. The capture of foreign genes and intragenomic sequence duplication might be driven by the same mechanism, i.e., copy-choice RNA recombination (Kirkegaard and Baltimore, 1986; Wang and Walker, 1993). On evolutionary divergence of closteroviruses, some of these elements underwent further shuffling. Thus the sequence coding for the leader P-Pro was duplicated in the CTV genome, and the gene for the CP homologue was extended and moved downstream of the (major) CP gene in the genomes of mono- and bipartite whitefly-borne closteroviruses (or vice versa). The N- and C-terminal domains of the CP homologue of LIYV both showed similarity to the CP, so it cannot be ruled out that a triplication of the CP gene has occurred in the LIYV genome (Klaassen *et al.,* 1995). Consistent with this hypothesis, the sizes of the CP duplicates of SPSVV (79 kDa) and CCSV (74 kDa), as well as that of LIYV (52 kDa), are rough multiples of their CP sizes (~28 kDa), and the repeated segments of marginal similarity to the CP core may also be found in the N-terminal parts of the 79- and 74-kDa proteins (author, unpublished observation). Thus, expansion of the closterovirus genomes may be partly attributed to insertions and tandem duplications at both ends of the conserved replicative core.

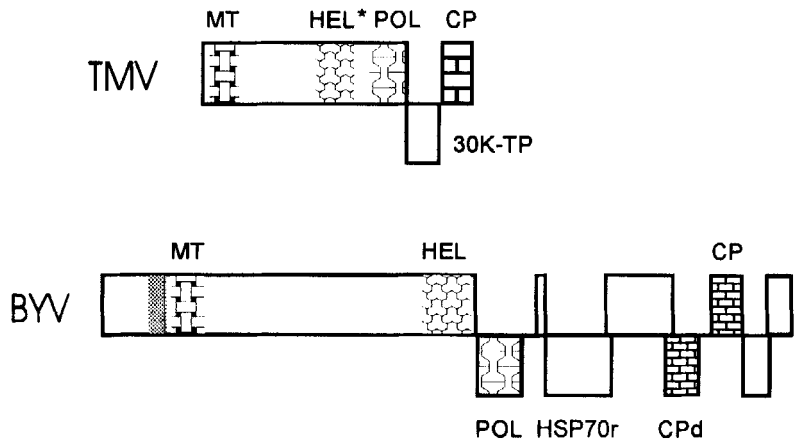

FIG. 6. Schematic comparison of the genomes of a closterovirus (BYV) and a tobamovirus (TMV). Related domains are shown by identical fill patterns and are marked as in Fig. 2. Asterisk marks the leaky termination codon in the TMV replicase gene; 30K, transport or movement protein gene of TMV; CP, capsid protein.

Another large insertion in the closterovirus genomes lies within the replicative core, between the MT and HEL domains (Fig. 6). The sequence of this region is significantly diverged in the BYV, CTV, and LIYV replicases (Dolja *et al.*, 1994; Karasev *et al.*, 1995; Klaassen *et al.*, 1995). Interestingly, among closteroviruses and related plant viruses the size of the MT–HEL span grows almost linearly with the increase of the genome size; in the case of viruses with divided genomes, it is related rather to the size of the largest genomic component (Table III). Generally, a rule "the larger the genome, the larger the replicase" inferred from these comparisons may also be applicable to arteri-, toro-, and coronaviruses of animals (den Boon *et al.*, 1991; Godeny *et al.*, 1993; Snijder and Horzinek, 1993). This relationship is not trivial, as the overall increase in genome size in these cases is not solely due to insertions in the replicase gene(s), but also to the appearance of new coding sequences flanking the replicative core. It seems quite likely that, in the process of evolution, expansion of closterovirus genomes was attended by an increase in the size of their replicases (Agranovsky, *et al.*, 1994a). At least one apparent obstacle in maintaining large RNA genomes must be accumulation of mutations in the progeny strands due to low fidelity of viral RNA polymerases (Holland *et al.*, 1982; Steinhauer and Holland, 1987). It would be interesting to see if, for example, the inserted domains or the leader proteins serve as "spell-

TABLE III

RELATIONSHIPS BETWEEN THE GENOME LENGTH AND REPLICASE SIZE IN THE TOBAMO-LIKE LINEAGE OF PLANT VIRUSES

Virus (group)	Genome size (kb)	Replicase size (aa)	MT–HEL distance (aa)[a]	HEL–POL distance (aa)[b]
BMV (bromo)	3.2+2.9+2.1	960+820[c]	430	
BSMV (hordei)	3.8+3.3+3.2	1140+770[c]	540	
RBDV (idaeo)	5.5+2.2	1690	420	680
TMV (tobamo)	6.4	1620	560	640
TRV (tobra)	6.8+3.4	1710	610	640
SBWMV (furo)	7.1+3.6	1830	700	650
LIYV (clostero)	8.1+7.2	2070[d]	1000	660
BYV (clostero)	15.5	2510[d]	1400	660
CTV (clostero)	19.3	2610[d]	1470	660

[a] The MT–HEL distance is measured between the C-terminal part of methyltransferase motif IV (Rozanov *et al.*, 1992) and the GKS/T signature in helicase motif I (Gorbalenya and Koonin, 1993).

[b] The HEL–POL distance is measured between the GKS/T and the GDD polymerase signature (Kamer and Argos, 1984).

[c] For BMV and BSMV, in which MT–HEL and POL domains are assigned to two individual proteins, sizes are given for both the putative methyltransferase-helicase and polymerase.

[d] The sizes of closterovirus replicases are given assuming translational frameshifting for ORFs 1a/1b expression and the cleavage of the N-terminal leader protein.

checkers" on strand copying. In DNA-dependent DNA polymerases, the 3′–5′ exonuclease activity assigned to a distinct protein domain is crucial for high replication fidelity (reviewed in Kunkel, 1988). Also, these domains might mediate homologous recombination between the virus RNA molecules to get rid of incorrigible errors, thus maintaining viable progeny. In coronavirus replication, recombination is believed to be a key mechanism to combat high-frequency errors (Jarvis and Kirkegaard, 1991; Lai, 1990). Finally, there is a possibility that the insert between MT and HEL contains a set of distinct domains to recognize the replication signals on an RNA template. The fact that up to six and nine sgRNA species may be synthesized on BYV and CTV infection, respectively, as compared with only two sgRNAs in the case of TMV, may again be a corollary to the increased complexity of the closterovirus replicases. Naturally, the possibilities mentioned above do not exclude one another.

In sharp contrast to the MT–HEL span, the distance between HEL and POL is essentially the same, about 650 residues, in all the tobamo-like virus replicases, despite the POL expression mode (Table III). Conservation of this arrangement may reflect constraints imposed on the replicase architectonics that must ensure concerted action of the strand-separating helicase "wedge" and the copying polymerase unit. The fact that the HEL and POL domains are found in two distinct gene products of bromo- and hordeiviruses does not contradict this rule as it would seem; at least for brome mosaic bromovirus, it has been demonstrated that the helicase-like and polymerase-like proteins form a complex in which the HEL and POL domains are juxtaposed in a fashion very similar to that in the TMV replicase (Kao and Ahlquist, 1992; Kao *et al.*, 1992).

B. Packaging Constraints and RNA Genome Evolution

Apart from replication constraints, evolution of closteroviruses toward increasing the genome size would have had to overcome packaging constraints. Comparisons of particle and genome structure of closteroviruses with those of other plant RNA viruses reveal some tendencies that may help one imagine how this could happen. Mono- and multipartite RNA viruses can be subdivided into those having "compressed" and "stretched" genomes, and this may be related to the virion type. Some spherical viruses, namely luteoviruses, tombusviruses, and tymoviruses, have compressed monopartite genomes in which ORFs extensively overlap to form "doubledecker" gene arrangements (Miller *et al.*, 1988; Morch *et al.*, 1988; Rochon and Tremaine, 1989). Such economical use of the coding sequence perhaps reflects a compromise between the necessity to widen the repertoire of viral genes and the limited size of an RNA molecule that would fit a spherical particle (Bransom *et al.*, 1995). However, the maintenance of overlapping genes has an apparent drawback, as this precludes each gene from being optimally adapted (Keese and Gibbs, 1992). The genome splitting seen in many RNA virus groups perhaps allows lifting of packaging constraints and minimizing the use of overlapping ORFs (or decompressing the preexisting gene overlaps). Thus, in spherical comoviruses, nepoviruses, dianthoviruses, and *Bromoviridae* members, the genomes are divided among separately encapsidated RNA components, each containing nonoverlapping gene(s) (Fig. 7).

An elongated helical capsid is less restrictive for the size of enveloped RNA. As proposed for the corona-like viruses, transition from a spherical (arterivirus-type) to helical (coronavirus-type) nucleocapsid

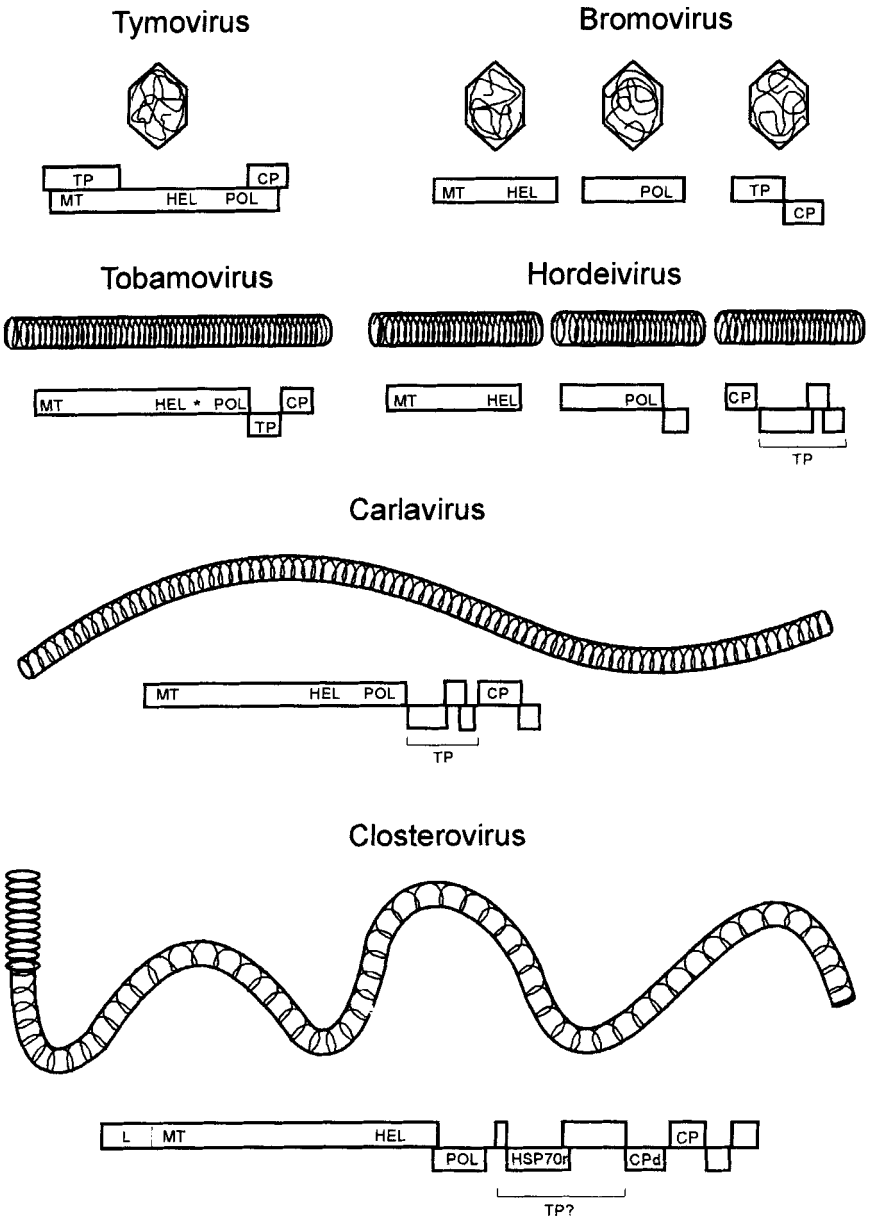

FIG. 7. Relationships between capsid type and genome organization in closteroviruses and other Sindbis-like supergroup viruses. Designations of the conserved protein domains are as in Fig. 2; other designations are TP, transport or movement protein(s); L, leader protein, *, leaky stop codon. Depicted roughly to scale.

exempted a progenitor of the toro- and coronaviruses from packaging constraints, thus allowing a nearly twofold genome expansion (Godeny *et al.*, 1993). Characteristically, most of the plant RNA viruses with elongated particles (tobamo-, tobra-, furo-, poty-, carla-, and closteroviruses) have stretched genomes, with modest overlapping of genes if any (Fig. 7). However, a subset of these viruses having rigid rodlike particles do not encapsidate RNA molecules of more than approximately 7 kb, possibly because of restrictions imposed on the particle length by sterical hindrances in the cell (Dolja *et al.*, 1994) and/or their mechanical fragility. This could have forced the splitting of the genomes of tobra-, hordei-, and furoviruses. In contrast to tobamoviruses having a monopartite 6.4-kb genome with only three genes, these viruses have genomes of 9–10 kb that encompass four to seven genes (Fig. 7).

It has been reasonably hypothesized that acquisition of flexible and superflexible helical capsids by the ancestors of carlaviruses, potyviruses, and closteroviruses allowed their genomes to grow to 10 kb and 20 kb, respectively (Dolja *et al.*, 1994). Again, these viruses possess stretched monopartite genomes (Fig. 7). As for the possible relationship between genome division and capsid type, the existence of bipartite filamentous viruses allied with the last two groups (bymoviruses and bicomponent closteroviruses) suggests the involvement of evolutionary factors other than packaging constraints that might have driven the genome splitting. In closteroviruses, the capsid evolution was crowned by employing the second CP. This conferred on their particles a structural complexity unprecedented among simple elongated viruses, which may be expected to require unusual assembly mechanisms.

VIII. Conclusions

The borrowing from Boris Pasternak's book of verse ("Over the Barriers," 1914–1916) in the title emphasizes that closteroviruses evolved by surmounting the restraints imposed on the genome and particle structure of positive-strand RNA viruses. Closteroviruses have large RNA genomes whose size and coding potential may only be compared to those of the corona-like viruses. Nevertheless, despite similar expression strategies and genome layouts developed in these two groups, closteroviruses cannot be considered as "plant coronaviruses" of a kind, since these similarities do not extend to amino acid sequences. Rather, they reflect independent adaptation to handling large RNA genomes in the two evolutionarily distant lineages. Colinearity and conservation of the main replicative domains clearly suggest the common ancestry of

closteroviruses and other plant tobamo-like viruses. However, closteroviruses have followed a distinct evolutionary pathway that has led to dramatic expansion of their genomes. Along with this, their evolution would have had to solve problems connected with replication and packaging of large RNA molecules; it is plausibe that this has been achieved by increasing the size (and functional complexity) of RNA replicase and by using a superflexible capsid made up of two CPs. Expansion of the closterovirus genomes has partially resulted from RNA recombination. It is possible that the horizontally acquired elements brought in novel enzymatic activities and structural elements advantageous for closterovirus adaptation to a distinct ecological niche, distinguished by the phloem-limited nature of infection and the semipersistent mode of insect transmission. In this respect, the most intriguing products are the HSP70-related protein, having the properties of a microtubule-binding ATPase, and the capsid protein duplicate involved in particle formation. There are many more closterovirus gene products whose functions remain enigmatic, since they have neither sequence-related counterparts in a current database nor known functional motifs. Hence, we may have more surprises. At present, we are making only the first steps in perceiving how closterovirus infection proceeds at the molecular level, despite some progress that has made it possible at least to address these questions. Further studies of the functions encoded by the large RNA genomes of closteroviruses are expected to provide a better understanding of the molecular mechanisms of their interactions with the genomes of their hosts and vectors.

ACKNOWLEDGMENTS

I am especially grateful to U. Hoyer, E. Maiss, and H. J. Vetten; to W. Jelkmann; and to L. P. Woudt and S&G Seeds B.V. for communicating their sequence data prior to publication. I wish to thank V. P. Boyko, K. Eastwell, B. W. Falk, S. Gowda, and J. P. Martelli for providing their unpublished data and preprints, D.-E. Lesemann for electron micrographs, and F. Rippmann for communicating his protein modelbuilds. Thanks are due to J. G. Atabekov, Yu. L. Dorokhov, A. V. Galkin, A. E. Gorbalenya, E. V. Koonin, and S. Yu. Morozov for fruitful discussions and critical reading of the manuscript.

REFERENCES

Agranovsky, A. A., Boyko, V. P., Karasev, A. V., Koonin, E. V., and Dolja, V. V. (1991a). *J. Mol. Biol.* **217**, 603–610.

Agranovsky, A. A., Boyko, Y. P., Karasev, A. V., Lunina, N. A., Koonin, E. V., and Dolja, V. V. (1991b). *J. Gen. Virol.* **72**, 15–23.

Agranovsky, A. A., Koonin, E. V., Boyko. V. P., Maiss, E., Frötschl, R., Lunina, N. A., and Atabekov, J. G. (1994a). *Virology* **198**, 311–324.

Agranovsky, A. A., Koenig, R., Maiss, E., Boyko, V. P., Casper, R., and Atabekov, J. G. (1994b). *J. Gen. Virology* **75**, 1431–1439.

Agranovsky, A. A., Lesemann, D.-E., Maiss, E., Hull, R., and Atabekov, J. G. (1995). *Proc. Natl. Acad. Sci. U.S.A.* **92**, 2470–2473.

Ahlquist, P., Strauss, E. G., Rice, C. M., Strauss, J. M., Haseloff, J., Zimmern, D. (1985). *J. Virol.* **53**, 536–542.

Atreya, C. D., Atreya, P. L., Thornbury, D. W., and Pirone, T. P. (1992). *Virology* **191**, 106–111.

Bahner, I., Lamb, J., Mayo, M. A., and Hay, R. T. (1990). *J. Gen. Virol.* **71**, 2251–2256.

Bar-Joseph, M., and Hull, R. (1974). *Virology* **62**, 552–562.

Bar-Joseph, M., and Lee, R. F. (1989). CMI/AAB Descript. Plant Viruses. No. 353.

Bar-Joseph, M., and Murant, A. F. (1982). CMI/AAB Descript. Plant Viruses. No. 260.

Bar-Joseph, M., Loebenstein, G., and Cohen, J. (1972). *Virology* **50**, 821–828.

Bar-Joseph, M., Inouye, T., and Sutton, J. (1976). *Plant Dis. Rep.* **60**, 851.

Bar-Joseph, M., Garnsey, S. M., and Gonsalves, D. (1979). *Adv. Virus Res.* **25**, 93–168.

Beck, D. L., Guilford, P. J., Voot, D. M., Andersen, M. T., and Forster, R. L. S. (1991). *Virology* **183**, 695–702.

Belli, G., Fortusini, A., Casati, P., Belli, L., Bianco, P. A., and Pratti, S. (1994). *Riv. Pat. Veg. S.V.* **4**, 105–108.

Bem, F., and Murant, A. F. (1979). *Ann. Appl. Biol.* **92**, 237–242.

Boscia, D., Greif, C., Gugerli, P., Martelli, G. P., Walter, B., and Gonsalves, D. (1995). *VITIS* **34**, in press.

Boyko, V. P., Karasev, A. V., Agranovsky, A. A., Koonin, E. V., and Dolja, V. V. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9156–9160.

Brandes, J., and Bercks, R. (1965). *Adv. Virus Res.* **11**, 1–12.

Bransom, K. L., Weiland, J. J., Tsai, C.-H., and Dreher, T. W. (1995). *Virology* **206**, 403–412.

Brault, V., and Miller, W. A. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2262–2266.

Brierly, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Block, V. C., Brown, T. D. K., and Inglis, S. C. (1987). *EMBO J.* **6**, 3779–3785.

Brunstedt, J., Moseley, J., and Hull, R. (1991). *Virus Genes* **5**, 267–272.

Butler, P. J. G. (1984). *J. Gen. Virol.* **65**, 253–279.

Carrington, J. C., Cary, S. M., Parks, T. D., and Dougherty, W. G. (1989). *EMBO J.* **8**, 365–370.

Cheng, S.-L., Domier, L. L., and D'Arcy, C. (1994). *Virology* **202**, 1003–1006.

Chevallier, D., Engle, A., Wurtz, M., and Charles, P. (1983). *J. Gen. Virol.* **64**, 2289–2293.

Choi, G. H., Pawlyk, D. M., and Nuss, D. L. (1991). *Virology* **183**, 747–752.

Clare, J. J., Belcourt, M., and Farabaugh, P. J. (1988). *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6816–6820.

Coffin, R. S., and Coutts, R. H. A. (1992). *Intervirology* **33**, 197–203.

Coffin, R. S., and Coutts, R. H. A. (1993). *J. Gen. Virol.* **74**, 1475–1483.

Cohen, J., Franck, A., Yetten, H. J., Lesemann, D.-E., and Loebenstein, G. (1992). *Ann. Appl. Biol.* **121**, 257–268.

Craig, E. A., Gambill, B. D., and Nelson, R. J. (1993). *Microbiol. Rev.* **82**, 3616–3620.

Craigen, W. J., Cook, R. G., Tate, W. P., and Caskey, C. T. (1985). *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3616–3620.

Cronin, S., Verchot, J., Haldeman-Cahill, R. Schaad, M. G., and Carrington, J. C. (1995). *Plant Cell* **7**, 549–559.

den Boon, J. A., Snijder, E. J., Chirnside, E. D., de Vries, A. A. F., Horzinek, M. C., and Spaan, W. J. M. (1991). *J. Virol.* **65**, 2910–2920.

Dinman, J. D., Icho, T., and Wickner, R. B. (1991). *Proc. Natl. Acad. Sci. U.S.A.* **88**, 174–178.
Dobrov, E. V., and Atabekov, J. O. (1989). *In* "Plant Viruses," Vol. I, "Structure and Replication" (C. L. Mandahar, ed.), pp. 174–205. CRC Press, Boca Raton, FL.
Dodds, J. A., and Bar-Joseph, M. (1983). *Phytopathology* **73**, 419–423.
Dodds, J. A., Jarupat, T., Lee, J. G., and Roistacher, C. N. (1987). *Phytopathology* **77**, 442–447.
Dolja, V. V., Karasev, A. V., and Agranovsky, A. A. (1990). *In* "New Aspects of Positive-Strand RNA Viruses" (R. Rueckert and M. Brinton, eds.), pp. 31–35. ASM Publications, Washington, DC.
Dolja, V. V., Boyko, V. P., Agranovsky, A. A., and Koonin, E. V. (1991). *Virology* **184**, 79–86.
Dolja, V. V., Karasev, A. V., and Koonin, E. V. (1994). *Annu. Rev. Phytopathol.* **32**, 261–285.
Domingo, E., Sabo, T., Taniguchi, T., and Weissmann, C. (1978). *Cell* **13**, 735–744.
Duffus, J. E. (1972). *Phytopathology* **62**, 3161–3165.
Duffus, J. E. (1973). *Adv. Virus Res.* **18**, 347–386.
Duffus, J. E., Larsen, R. C., and Liu, H. Y. (1986). *Phytopathology* **76**, 97–100.
Dunigan, D. D., and Zaitlin, M. (1990). *J. Biol. Chem.* **265**, 7779–7786.
Essau, K., and Hoefert, L. L. (1971). *Protoplasma* **72**, 255.
Falk, B. W., and Duffus, J. E. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 275–296. Plenum Press, New York.
Farabaugh, P. J. (1993). *Cell* **74**, 591–596.
Farabaugh, P. J., Zhao, H., and Vimaladithan, A. (1993). *Cell* **74**, 93–103.
Febres, V. J., Pappu, H. R., Anderson, E. J., Pappu, S. S., Lee, R. F., and Niblett, C. L. (1994). *Virology* **201**, 178–181.
Felsenstein, J. (1989). "PHYLIP 3.2 Manual." Herbarium, University of California, Berkeley.
Filichkin, S. A., Lister, R. M., McGrath, P. F., and Young, M. J. (1994). *Virology* **205**, 290–299.
Francki, R. I. B., Faquet, C. M., Knudson, D. L., and Brown, F., eds. (1991). *Arch. Virol.* **Suppl. 2**, 1.
German, S. Candresse, T., Lanneau, M., Huet, J. C., Pernollet, J. C., and Dunez, J. (1990). *Virology* **179**, 104–112.
Gething, M. J., and Sambrook, J. (1992). *Nature* **355**, 33–45.
Gibbs, A. (1987). *J. Cell Sci.* (Suppl.) **7**, 319–337.
Godeny, E. K., Chen, L., Kumar, S., Methven, S. L., Koonin, E. V., and Brinton, M. A. (1993). *Virology* **194**, 585–596.
Goldbach, R., Le Gall, O., and Wellink, J. (1991). *Semin. Virol.* **2**, 19–25.
Gorbalenya, A. E. (1992). *Semin. Virol.* **3**, 359–371.
Gorbalenya, A. E., and Koonin, E. V. (1989). *Nucleic Acids Res.* **17**, 8413–8440.
Gorbalenya, A. E., and Koonin, E. V. (1993). *Curr. Opin. Struct. Biol.* **3**, 419–429.
Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M. (1988). *Nature* **333**, 22.
Gorbalenya, A. E., Blinov, V. M., Donchenko, A. P., and Koonin, E. V. (1989). *J. Mol. Evol.* **28**, 256–268.
Guerri, J., Moreno, P., Munoz, N., and Martinez, M. E. (1991). *Plant Pathol.* **40**, 38–44.
Gunasinghe, U. B., and German, T. L. (1989). *Phytopathology* **79**, 1337–1341.
Haeberle, A.-M., Stussi-Garaud, C., Schmitt, C., Garaud, J.-C., Richards, K. E., Guilley, H., and Jonard, G. (1994). *Arch. Virol.* **134**, 195–203.
Hilf, M. E., Karasev, A. V., Pappu, H. R., Gumpf, D. J., Niblett, C. L., and Garnsey, S. M. (1995). *Virology* **208**, 576–582.
Hill, H. R., and Zetler, F. W. (1973). *Phytopathology* **63**, 443.

Hills, G. J., and Gay, M. R. (1976). *John Innes Inst. Annu. Report* **67**, 107–109.
Hodgman, T. C. (1988). *Nature* **333**, 22–23; 578 (Erratum).
Holland, J. J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and Van de Pol, S. (1982). *Science* **215**, 1577–1585.
Hu, J. S., Gonsalves, D., and Tellis, O. (1990). *J. Phytopath. Z.* **128**, 1–14.
Hull, R. (1992). *Semin. Virol.* **3**, 373–382.
Inouye, T. (1974). CMI/AAB Descript. Plant Viruses. No. 136.
Inouye, T. (1976). CMI/AAB Descript. Plant Viruses. No. 157.
Jacks, T., and Varmus, H. E. (1985). *Science* **230**, 1237–1242.
Jacks, T., Madhani, H. D., Masiarz, F. R., and Varmus, H. E. (1988). *Cell* **55**, 447–458.
Jarvis, T. C., and Kirkegaard, K. (1991). *Trends Genet.* **7**, 186–191.
Jiang, B. M., Monroe, S. S., Koonin, E. V., Stine, S. E., and Glass, R. I. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 10539–10543.
Kamer, G., and Argos, P. (1984). *Nucleic Acids Res.* **12**, 7269–7283.
Kao, C. C., and Ahlquist, P. (1992). *J. Virol.* **66**, 7293–7302.
Kao, C. C., Quadt, R., Hershberger, R. P., and Ahlquist, P. (1992). *J. Virol.* **66**, 7293–7302.
Karasev, A. V., Agranovsky, A. A., Rogov, V. V., Miroshnichenko, N. A., Dolja, V. V., and Atabekov, J. G. (1989). *J. Gen. Virol.* **70**, 241–245.
Karasev, A. V., Kashina, A. S., Gelfand, V. I., and Dolja, V. V. (1992). *FEBS Lett.* **304**, 12–14.
Karasev, A. V., Nikolaeva, O. N., Gumpf, D. J., Garnsey, S. M., and Dawson, W. O. (1994a). *Phytopathology* **84**, 1156.
Karasev, A. V., Nikolaeva, O. N., Koonin, E. V., Gumpf, D. J., and Garnsey, S. M. (1994b). *J. Gen. Virol.* **75**, 1415–1422.
Karasev, A. V., Boyko, V. P., Gowda, S., Nikolaeva, O. N., Hilf, M. E., Koonin, E. V., Niblett, C .L., Cline, K. C., Gumpf, R. F., Gamsey, S. M., Lewandowski, D. J., and Dawson, W. O. (1995). *Virology* **208**, 511–520.
Keese, P. K., and Gibbs, A. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 9489–9493.
Kim, K. S., Gonsalves, D., Teliz, D., and Lee, K. W. (1989). *Phytopathology* **79**, 357–360.
Kirkegaard, K., and Baltimore, D. (1986). *Cell* **47**, 433–443.
Klaassen, V. A., Boeshore, M. L., Dolja, V. V., and Falk, B. W. (1994). *J. Gen. Virol.* **75**, 1525–1533.
Klaassen, V. A., Boeshore, M. L., Koonin, E. V., Tian, T., and Falk, B. W. (1995). *Virology* **208**, 99–101.
Koenig, R., and Torrance, L. (1986). *J. Gen. Virol.* **67**, 2145–2151.
Koonin, E. V. (1991). *J. Gen. Virol.* **72**, 2197–2206.
Koonin, E. V., and Dolja, V. V. (1993). *Crit. Rev. Biochem. Mol. Biol.* **28**, 375–430.
Koonin, E. V., Mushegian, A. R., Ryabov, E. V., and Dolja, V. V. (1991). *J. Gen. Virol.* **72**, 2895–2903.
Kunkel, T. A. (1988). *Cell* **53**, 837–840.
Lai, M. M. C. (1990). *Annu. Rev. Microbiol.* **44**, 303–333.
Lain, S., Riechmann, J. L., and Garcia, J. A. (1990). *Nucleic Acids Res.* **18**, 7003–7006.
Larsen, R. S., Kim, K. S., and Scott, H. A. (1991). *Phytopathology* **81**, 227–232.
Lee, H.-J., Shieh, C.-K., Gorbalenya, A. E., Koonin, E. V., La Monica, N., Tuler, J., Bagdzhadzhyan, A., and Lai, M. M. C. (1991). *Virology* **180**, 567–582.
Lee, H.-J., Calvert, L. A., Nagel, J., Hubbard, J. D. (1988). *Phytopathotogy* **78**, 1221–1226.
Lesemann, D.-E. (1977). *Phytopathol. Z.* **89**, 330–334.
Lesemann, D.-E. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 179–236. Plenum Press, New York.
Limburg, D. D., Mauk, P. A., and Godfrey, L. D. (1994). *Phytopathology* **84**, 1066.

Ling, K.-S., Drong, R. F., Slightom, J. L., and Gonsalves, D. (1994). *Phytopathology* **84**, 1372.

Lister, R. M., and Bar-Joseph, M. (1981). *In* "Handbook of Plant Virus Infections and Comparative Diagnosis" (E. Kurstak, ed.), pp. 809–844. Elsevier/North-Holland, Amsterdam.

Liu, H. Y., and Duffus, J. E. (1990). *Phytopathology* **80**, 866–869.

Lockhart, B. E. L., Autrey, L. J. C., and Comstock, J. C. (1992). *Phytopathology* **82**, 691–695.

Lomonossoff, G. P., and Wilson, T. M. A. (1985). *In* "Molecular Plant Pathology" (J. W. Davies, ed.), Vol. 1, pp. 43–83. CRC Press, Boca Raton, FL.

Mäkinen, K., Næss, V., Tamm, T., Truve, E., Aaspõlu, A., and Saarma, M. (1995). *Virology* **207**, 566–571.

Mathews, R. E. F. (1991). "Plant Virology" (3rd Ed.). Academic Press, New York.

Mawassi, M., Gafny, R., Gagliardi, D., and Bar-Joseph, M. (1995a). *J. Gen. Virol.* **76**, 651–659.

Mawassi, M., Karasev, A. V. , Mietkiewska, E., Gafny, R., Lee, R. F., Dawson, W. O., and Bar-Joseph, M. (1995b). *Virology* **208**, 383–387.

Mayo, M. A., and Martelli, G. P. (1993). *Arch. Virol.* **133**, 496–498.

Mi, S., and Stollar, V. (1991). *Virology* **178**, 429–434.

Miller, W. A., Waterhouse, P. M., and Gerlach, W. L. (1988). *Nucleic Acids Res.* **16**, 6097–6111.

Milne, R. G. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4. Plenum Press, New York.

Morch, M.-D., Boyer, J.-C., and Haenni, A.-L. (1988). *Nucleic Acids Res.* **16**, 6157–6173.

Moreno, P., Guero, J., and Munoz, N. (1990). *Phytopathology* **80**, 477–482.

Morozov, S. Yu., Dolja, V. V., and Atabekov, J. G. (1989). *J. Mol. Evol.* **297**, 52–62.

Morozov, S. Yu., Miroshnichenko, N. A., Zelenina, D. A., Fedorkin, O, N., Soloviev, A. G., Lukasheva, L. I., and Atabekov, J. G. (1990). *Biochemie* **72**, 677–684.

Moseley, J., and Hull, R. (1990). *Ann. Appl. Biol.* **118**, 605–613.

Murant, A .F., Raccah, B., and Pirone, T. P. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 237–274. Plenum Press, New York.

Nakano, M., and Inouye, T. (1980). *Ann. Phytopath. Soc. Jpn.* **46**, 7–10.

Namba, S., Boscia, D., Azzam, O., Maixner, M., Hu, J. S., Golino, D., and Gonsalves, D. (1991). *Phytopathology* **81**, 964–970.

Nikiphorova, S. Yu., Agranovsky, A. A., and Atabekov, J. C. (1995). *Dokladi Rossiiskoi Akademii Nauk* **340**, 416–418 (in Russian).

Pappu, H. R., Pappu, S. S., Manjunath, K. L., Lee, R. F., and Niblett, C. L. (1993). *Proc. Natl. Acad. Sci. U.S.A.* **90**, 3641–3644.

Pappu, H. R., Karasev, A. V., Anderson, E. J., Pappu, S. S., Hilf, M. E, Febres, M. E., Eckloff, R. M. G., McCaffery, M., Boyko, V., Gowda, S., Dolja, V. V., Koonin, E. V., Gumpf, D. J., Cline, K. C., Garnsey, S. M., Dawson, W. O., Lee, R. F., and Niblett, C. L. (1994). *Virology* **199**, 35–46.

Pirone, T. P. (1991). *Semin. Virol.* **2**, 81–87.

Prüfer, D, Tacke, E., Schimtz, J., Kull, B., Kaufmann, A., and Rohde, W. (1992). *EMBO J.* **11**, 1111–1117.

Ragetti, H. W. J., Elder, M., and Schroeder, B. K. (1982). *Can. J. Bot.* **60**, 1235–1248.

Raine, J., McMullen, R. D., and Forbes, A. R. (1986). *Can. J. Plant Pathol.* **8**, 6–11.

Richards, K. E., and Tamada, T. (1992). *Annu. Rev. Phytopathol.* **30**, 291–313.

Rippmann, F., Taylor, W. R., Rothbard, J. B., and Green, M. N. (1991). *EMBO J.* **10**, 1053–1059.

Rochon, D. M., and Tremaine, J. H. (1989). *Virology* **169**, 251–259.
Rogov, V. V., Karasev, A. A., and Agranovsky, A. A. (1993). *Phytopathol. Z.* **137**, 79–88.
Rozanov, M. N., Koonin, E. V., and Gorbalenya, A. E. (1992). *J. Gen. Virol.* **73**, 2129–2134.
Russel, M. (1993). *J. Mol. Biol.* **231**, 689–697.
Sekiya, M. E., Lawrence, S. D., McCaffery, M., and Cline, K. (1991). *J. Gen. Virol.* **72**, 1013–1020.
Schmidt, H. B., Richter, J., Hertsch, W., and Klinkowski, M. (1963). *Phytopathol. Z.* **47**, 66–74.
Shukla, D. D., Strike, P. M., Tracy, S. L., Gough, K. H., and Ward, C. W. (1988). *J. Gen. Virol.* **69**, 1497–1508.
Snijder, E. J., and Horzinek, M. C. (1993). *J. Gen. Virol.* **74**, 2305–2316.
Steinhauer, D. A., and Holland, J. J. (1987). *Annu. Rev. Microbiol.* **41**, 409–433.
Strauss, E. G., and Strauss, J. H. (1986). *In* "The Togaviridae and Flaviviridae" (E.G. Strauss and J. H. Strauss, eds.), pp. 35–90. Plenum Press, New York.
Stuart, K. D., Weeks, R., Gilbride, L., and Myler, P. J. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8596–8600.
ten Dam, E. B. (1995). Ph.D. Thesis, Leiden University, pp. 93–106.
ten Dam, E. B., Pleij, C. W., and Bosch, L. (1990). *Virus Genes* **4**, 121–136.
Tollin, P., and Wilson, H. R. (1988). *In* "The Plant Viruses. The Filamentous Plant Viruses" (R. G. Milne, ed.), Vol. 4, pp. 51–83. Plenum Press, New York.
Tollin, P., Wilson, H. R., Roberts, I. M., and Murant, A. F. (1992). *J. Gen. Virol.* **75**, 3045–3048.
Tu, C., Tzeng, T.-H., and Bruenn, J. A. (1992). *Proc. Natl. Acad. Sci. U.S.A.* **89**, 8636–8640.
Wang, Y., and Walker, P. J. (1993). *Virology* **195**, 719–731.
Warrener, P., and Collett, M. S. (1995). *J. Virol.* **69**, 1720–1726.
Weiss, R. B. (1984). *Proc. Natl. Accad. Sci. U.S.A.* **81**, 5797–5801.
Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F., and Gesteland, R. F. (1988). *EMBO J.* **7**, 1503–1507.
Winter, S., Purac, A., Legget, F., Frison, E. A., Rossel, H. W., and Hamilton, R. I. (1992). *Phytopathology* **82**, 869–875.
Woudt, L. P., de Rover, A. P., and Van Grinsven, M. Q. J. M. (1993a). "Proc. Int. Congr Plant Pathol., 6th, Montreal," p. 313.
Woudt, L. P., de Rover, A. P., de Haan, P. T., and Van Grinsven, M. Q. J. M. (1993b). "Int. Congr. Virol, 9th, Glasgow," p. 326.
Xiong, Z., and Lommel, S. A. (1989). *Virology* **171**, 543–554.
Yamashita, S., Ohki, S. T., Doi, Y., and Yora, K. (1976). *Ann. Phytopathol. Soc. Jpn.* **42**, 382–390.
Yoshikawa, N., Sasaki, E., Kato, M., and Takahashi, T. (1992). *Virology* **191**, 98–105.
Zaccomer, B., Haenni, A.-L., and Macaya, G. (1995). *J. Gen. Virol.* **76**, 231–247.
Zimmern, D. (1988). *In* "RNA Genetics" (J. J. Holland, E. R. Domingo, and P. Ahlquist, eds.), pp. 211–240. CRC Press, Boca Raton, FL.
Zimmerman, D., Bass, P., Legin, R., and Walter, B. (1990). *J. Phytopathol. Z.* **130**, 205–218.