

Probable Reassortment of Genomic Elements among Elongated RNA-Containing Plant Viruses

Sergei Yu. Morozov, Valerian V. Dolja, and Joseph G. Atabekov

Department of Virology, Moscow State University, Moscow 119899, USSR

Summary. The relationships of genome organization among elongated (rod-shaped and filamentous) plant viruses have been analyzed. Sequences in coding and noncoding regions of barley stripe mosaic virus (BSMV) RNAs 1, 2, and 3 were compared with those of the monopartite RNA genomes of potato virus X (PVX), white clover mosaic virus (WCIMV), and tobacco mosaic virus, the bipartite genome of tobacco rattle virus (TRV), the quadripartite genome of beet necrotic yellow vein virus (BNYVV), and icosahedral tricornaviruses. These plant viruses belong to a supergroup having 5'-capped genomic RNAs. The results suggest that the genomic elements in each BSMV RNA are phylogenetically related to those of different plant RNA viruses. RNA 1 resembles the corresponding RNA 1 of tricornaviruses. The putative proteins encoded in BSMV RNA 2 are related to the products of BNYVV RNA 2, PVX RNA, and WCIMV RNA. Amino acid sequence comparisons suggest that BSMV RNA 3 resembles TRV RNA 1. Also, it can be proposed that in the case of monopartite genomes, as a rule, every gene or block of genes retains phylogenetic relationships that are independent of adjacent genomic elements of the same RNA. Such differential evolution of individual elements of one and the same viral genome implies a prominent role for gene reassortment in the formation of viral genetic systems.

Key words: Genome organization — Evolution — Plant virus — RNA recombination — Sequence similarity

Introduction

Sequencing of the genomes of RNA-containing viruses has offered a wealth of data for comparing amino acid sequences of the proteins they encode. Regions of similarity have been found among virus-specific proteins of viruses heretofore held to be quite remote, such as tobacco mosaic virus (TMV, a tobamovirus), brome mosaic virus (BMV, a bromovirus), alfalfa mosaic virus (AIMV, an ilarvirus), and Sindbis virus (an alphavirus) (Cornelissen and Bol 1984; Haseloff et al. 1984; Kamer and Argos 1984). It has been proposed that these and related viruses be classified in one supergroup (Goldbach 1986). The second supergroup comprises plant como-, poty-, and nepoviruses and animal picornaviruses that also show extensive similarity in the amino acid sequences of a number of proteins and similar gene arrangements (Argos et al. 1984; Franssen et al. 1984; Allison et al. 1986; Goldbach 1986; Meyer et al. 1986; Domier et al. 1987). The most distinguishing structural feature of the genomic RNA supergroups is that the 5' end carries a cap in the first supergroup but a covalently linked protein in the second one.

The genome of any virus is a combination of coding sequences (genes) and noncoding regulatory sequences. The data available allow us to estimate the minimal number of functions (genes) essential for replication of (+)RNA plant viruses (in different viruses they can occur in different combinations): (1) a block of genes (usually not less than two) ensuring synthesis of virus-specific RNAs; (2) gene(s) responsible for the cell-to-cell transport of viral genetic material over the infected plant; and (3) one or more genes coding for the virus coat. Viruses

whose RNA is translated into a polyprotein must also have gene(s) for processing protease(s).

It is just this minimal set that is present in the TMV genome: genes for the 126-kd and 183-kd proteins involved in RNA replication, for the 30-kd "transport" protein, and for the coat protein (for review see Palukaitis and Zaitlin 1986). Functionally analogous genes are distributed among the three genomic components of bromo-, cucumo-, and ilarviruses (Symons 1985). The bipartite tobacco rattle virus (TRV, a tobavirus) possesses an additional gene coding for a cysteine-rich 16-kd protein of unknown function (Cornelissen et al. 1986). The quadripartite beet necrotic yellow vein virus (BNYVV, a furovirus) encodes no protein similar to the transport proteins of the above-mentioned viruses, but BNYVV RNA 2 contains additional open reading frames (ORFs) discussed below (Bouzoubaa et al. 1987; Lemaire et al. 1988).

The main objects of our analysis are two viruses whose genome structure has been studied for quite a few years by our group in cooperation with the groups of K.G. Skryabin and Y.V. Kozlov. These are barley stripe mosaic virus (BSMV), a type member of the hordeivirus group, and potato virus X (PVX), a type member of the potexvirus group. Our aim was to compare the genome organization of hordei- and potexviruses with that of some other plant RNA viruses.

Methods

Amino Acid Sequence Data. Amino acid sequence data were taken from the following papers: BSMV RNA 1 (Rupasov et al. 1986); BSMV RNA 2 (Gustafson and Armour 1986); BSMV RNA 3 (Rupasov et al. 1986; Gustafson et al. 1987); BNYVV RNA 1 (Bouzoubaa et al. 1987); BNYVV RNA 2 (Bouzoubaa et al. 1986); TMV (Goelet et al. 1982); TRV (Cornelissen et al. 1986; Hamilton et al. 1987); tobamoviral coat proteins (Altschuh et al. 1987); PVX (Morozov et al. 1987; Krayev et al. 1988); WCIMV (Forster et al. 1988); TYMV (Morch et al. 1988).

Sequence Alignments. Initial similarity searches of the viral noncoat proteins were made with the matrix comparison program DIAGON (Staden 1982) run on a Wicat-S computer. The scoring system for this comparison was based on the observed frequency of substitution of one amino acid for another in a number of protein families (Schwartz and Dayhoff 1978). Groups of amino acids were: A, S, T, P, and G; N, D, E, and Q; H, R, and K; M, L, I, and V; F, Y, and W. Detailed alignments were made by visual inspection.

Hydrophilicity Prediction. The sequences of the small noncoat proteins were analyzed for hydrophilicity using the method described by Hopp and Woods (1981). The averaged hydrophilicity values for heptapeptides from $n - 3$ to $n + 3$ ($n > 4$) were plotted against n . The prediction of membrane-bound properties of proteins was made by the methods of Eisenberg et al. (1984) and Mohana Rao and Argos (1986).

Construction of Phylogenetic Trees. Conserved sequences of putative RNA polymerases shown in Fig. 6A were aligned with

those of other (+)RNA-containing viruses according to Koonin et al. (1988). A similarity matrix depending on the pairwise alignment of sequences was calculated using the Dayhoff exchange weights (Schwartz and Dayhoff 1978; Koonin et al. 1988). All possible subtrees for four-point conditions were constructed by using the similarity matrix. The final rootless phylogenetic tree consists of the subtrees combined optimally according to the maximum topological similarity principle (Chumakov and Yushmanov 1988; Yushmanov and Chumakov 1988).

Results

Similarity between Tricornavirus RNA 1 and the Genomes of BSMV and PVX

Filamentous potexviruses have a monopartite 5'-capped (Sonenberg et al. 1978) and 3'-polyadenylated (Morozov et al. 1981) genome whose internal genes are expressed through formation of subgenomic RNAs (Guilford and Forster 1986; Dolja et al. 1987). By now the PVX and WCIMV genomes have been completely sequenced (Morozov et al. 1987; Forster et al. 1988; Krayev et al. 1988).

Hordeiviruses are the only group of rod-shaped viruses with tripartite genome (for review see Atabekov and Dolja 1986). BSMV RNAs 2 and 3 and a part of RNA 1 have already been sequenced (Afanasiev et al. 1986; Gustafson and Armour 1986; Rupasov et al. 1986; Gustafson et al. 1987). Together with bromo-, cucumo-, and ilarviruses, BSMV may be considered formally a tricornavirus. There are indeed certain similarities among the four groups. BSMV RNAs are capped and, like those of bromo- and cucumoviruses (but unlike those of ilarviruses, for review see Haenni et al. 1982), carry a 3'-terminal tyrosine-accepting tRNA-like structure (Agranovsky et al. 1981; Kozlov et al. 1984). BSMV RNA 1, like RNA 1 of tricornaviruses, codes for one large protein (Dolja et al. 1983b; Rupasov et al. 1986). This 120-kd protein of BSMV exhibits similarity (at least in the C-terminal region) with domain 1C of the RNA-replicating 109-kd protein of BMV (Quadt et al. 1988), with related proteins of other tricornaviruses, tobamoviruses, tobaviruses, and, to a lesser extent, potexviruses and furoviruses (Cornelissen and Bol 1984; Rupasov et al. 1986; Hamilton et al. 1987; Forster et al. 1988; Krayev et al. 1988) (Fig. 1). All these proteins belong to the first family of the viral NTP-motif-containing proteins related to the bacterial DNA-helicases (Gorbalenya et al. 1988a,b; Hodgman 1988).

However, BSMV differs markedly from tricornaviruses in the number and arrangement of other genomic elements. All BSMV RNAs contain an internal poly(A) tract (Agranovsky et al. 1982) directly following the UAA terminator codon of 3'-terminal genes (Gustafson and Armour 1986; Rupasov et al. 1986). BMV has such a tract only in RNA 3, and

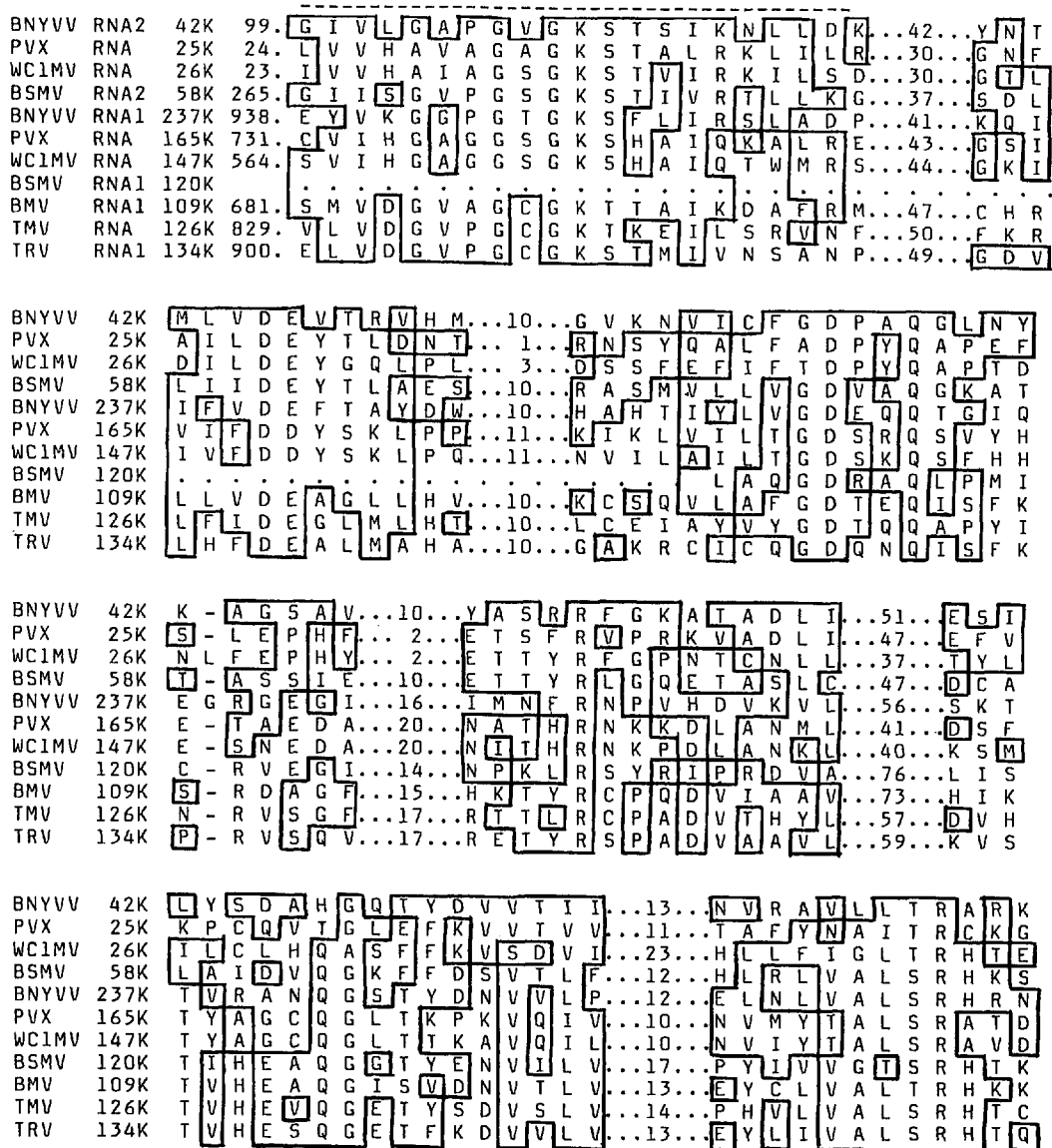


Fig. 1. Alignment of amino acid sequences of the 42-kd protein of BNYVV (Bouzoubaa et al. 1986), the 25-kd protein of PVX (Krayev et al. 1988), the 26-kd protein of WC1MV (Forster et al. 1988), and the 58-kd protein of BSMV (Gustafson and Armour 1986) with the putative proteins involved in replication of BNYVV (Bouzoubaa et al. 1987), PVX (Krayev et al. 1988), WC1MV (Forster et al. 1988), BSMV (Rupasov et al. 1986), BMV (Symons 1985), TMV (Palukaitis and Zaitlin 1986), and TRV (Hamilton et al. 1987). Positions where putative replicases have conserved similarities (Schwartz and Dayhoff 1978) with two or more sequences from the four top lines (including 42-kd BNYVV, 25-kd PVX, 26-kd WC1MV, 58-kd BSMV) are boxed. The amino acid sequence pattern similar to purine nucleotide-binding protein consensus is marked by a dashed underline. Numerals show the numbers of amino acid residues that could not be aligned. Gaps (-) have been introduced in the sequences for maximum alignment. The number in front of the sequences gives the position of its first amino acid in the protein.

there it is localized in the intergenic region (Ahlquist et al. 1981).

BSMV RNA 2: Triple Gene Block

On the other hand, in the array of genes, BSMV RNA 2 resembles RNA 2 of BNYVV. As can be seen in Fig. 2, both RNAs have the coat protein genes at the 5' termini—a situation rare enough among plant viruses. In BNYVV this is followed by

five other ORFs (Bouzoubaa et al. 1986, 1987) and in BSMV by three (Gustafson and Armour 1986). Comparison of amino acid sequences of the hypothetical proteins encoded therein shows that the possible products of BSMV RNA 2 have obvious, albeit variable, similarity with those of BNYVV RNA 2. Thus, the 14-kd and 17-kd proteins encoded in overlapping ORFs in the 3'-terminal region of BSMV RNA 2 have some properties of membrane-bound proteins as do also the 13-kd and 15-kd pro-

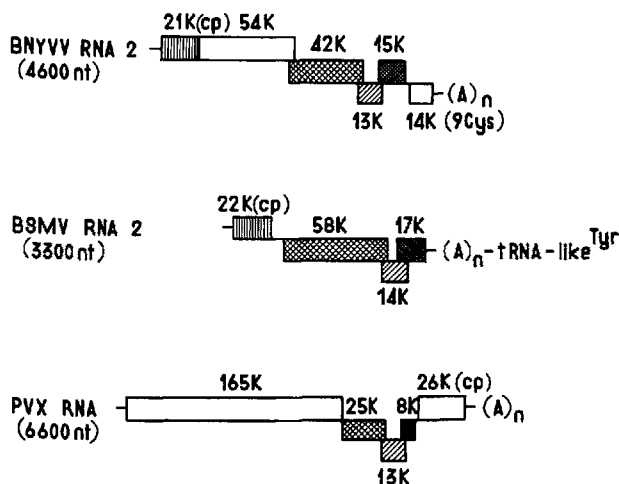


Fig. 2. Organization of BNYVV RNA 2 (adapted from Bouzoubaa et al. 1986), BSMV RNA 2 (from Gustafson and Armour 1986), and PVX RNA (Krayev et al. 1988). Counterparts are marked by identical shading.

teins of BNYVV (Morozov et al. 1987). These two pairs are similar in the relative location of hydrophobic and hydrophilic amino acid blocks (see Fig. 3). An analogous pair of proteins containing membrane-bound segments is encoded in overlapping ORFs (13 kd and 8 kd) of PVX. Moreover, BSMV 14-kd, BNYVV 13-kd, and PVX 13-kd proteins all carry a conserved amino acid block. This hydrophilic block separates two hydrophobic domains (Fig. 3).

The pairs of genes described are preceded by sequences coding for a 58-kd protein in BSMV RNA 2 (Gustafson and Armour 1986), a 42-kd protein in BNYVV RNA 2 (Bouzoubaa et al. 1986, 1987), and a 25-kd protein in PVX RNA (Fig. 2). As is evident from Fig. 1, these three proteins contain similar amino acid sequences. These sequences also display relation with the conserved domains of a family of predicted NTP-binding proteins of (+)RNA-containing plant viruses (Gorbalenya et al. 1985, 1988a; Dever et al. 1987). As suggested by Gorbalenya et al. (1988a), the above-mentioned proteins are the diverged relatives of the polypeptide domains encoded in the same genomic RNA of potexviruses (165-kd PVX and 147-kd WCIMV proteins), RNA 1 of BSMV (120-kd protein), and BNYVV (237-kd protein) (Figs. 1 and 2). Recently, Gorbalenya et al. (1988c) and Hodgman (1988) have shown that these viral proteins (belonging to the first family of NTP-motif-containing proteins) display a pronounced sequence similarity with the bacterial DNA helicases. It was speculated that these viral proteins may be involved in unwinding of double-stranded replicative forms during genomic RNA replication. In this case, predicted NTP-binding proteins encoded by the triple gene block could represent the subunits of

RNA helicases. Consequently, two NTP-binding proteins are encoded in the genomes of BSMV, BNYVV, PVX, and WCIMV. Thus, the genomes of three elongated plant viruses belonging to different groups (hordei-, furo-, and potexviruses) turn out to contain a similar block of three genes (Fig. 2), two of which probably code for membrane-bound proteins and the third for a predicted NTP-binding protein.

Comparison of the Capsid Proteins of BSMV and PVX with Other Viral Coat Proteins

Analysis of the amino acid sequences coded for by the 5'-terminal coat protein genes of BSMV and BNYVV reveals common motifs (Fig. 4). Similar motifs are found in the coat proteins of three other rod-shaped viruses: TMV, sunhemp mosaic virus (SHMV, a tobamovirus), and TRV. Interestingly, the corresponding genes in these latter viruses are located at the 3' end of TMV RNA (Fig. 5) and SHMV RNA (Meshi et al. 1981; Palukaitis and Zaitlin 1986) and on a separate nucleic acid, RNA 2, of TRV (Harrison and Robinson 1986).

The amino acid sequence of the PVX coat protein bears no appreciable similarity with those of BSMV, BNYVV, TMV, or TRV; it shows, however, obvious similarity with the coat protein sequences of a number of filamentous potyviruses (Morozov et al. 1987). It should be emphasized that despite similar particle morphology, potyviruses and potexviruses belong to different supergroups of RNA-containing viruses (see Introduction). In potyviruses the genome structure and mode of its expression are quite distinct from those of PVX, BSMV, or BNYVV (Allison et al. 1986; Domier et al. 1987).

BSMV RNA 3: Putative RNA-Polymerase Gene

The dicistronic RNA 3 of BSMV has yet no direct analogy among plant viruses in the number and arrangement of genes. Nevertheless, the proteins it codes for exhibit significant similarity with two out of four proteins encoded in RNA 1 of the bipartite TRV. Among putative polymerases of the plant Sindbis-like viruses (Goldbach 1986; Koonin et al. 1988), BSMV RNA polymerase, whose gene is located in the 5'-proximal part of RNA 3 (Rupasov et al. 1986; Gustafson et al. 1987), shows the greatest similarity with the putative TRV polymerase (the C-terminal region of the 194-kd protein encoded in RNA 1) (see Figs. 5 and 6A). This conclusion is confirmed illustratively by a phylogenetic rootless tree constructed using the aligned amino acid sequences shown in Fig. 6A and the sequences of other putative viral RNA polymerases (Fig. 6B). This dendrogram representing phylogenetic relationships

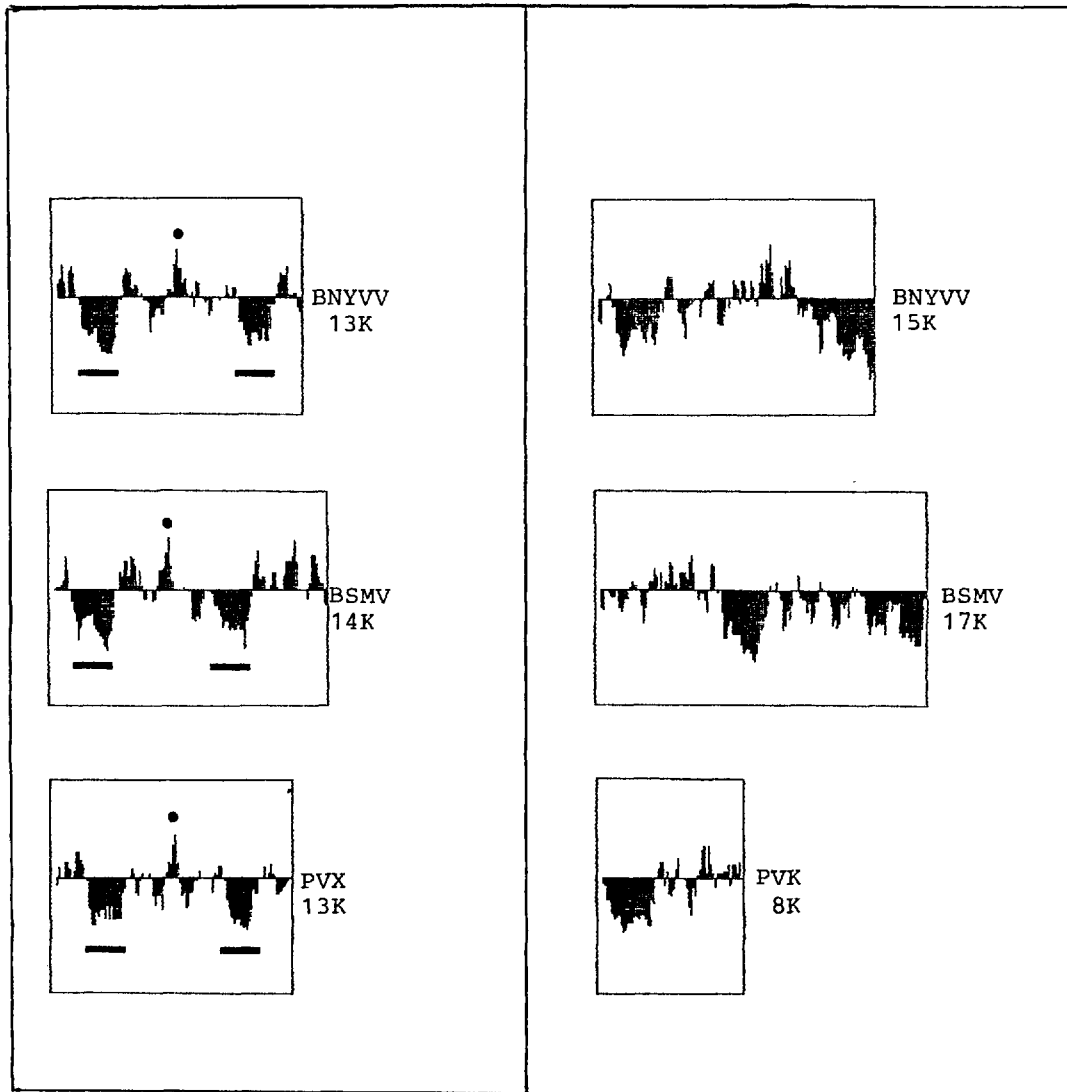


Fig. 3. Comparison of the hydrophilicity profiles of the putative membrane-bound proteins of BNYVV, BSMV, and PVX according to the method of Hopp and Woods (1981). Regions with a net hydrophobicity appear below the horizontal center line, and regions with a net hydrophilicity appear above the center line. Filled circles indicate positions for conserved hydrophilic amino acid blocks. Filled boxes are below positions for conserved hydrophobic segments (Morozov et al. 1987).

BSMV- 68.	D	R	D	L	P	A	D	V	D	R	R	F	A	G	A	R	G	A	R	G	Q	I	G	S	
TRV-PSG- 58.	S	R	Q	C	P	Y	S	P	Y	E	Q	F	P	E	R	K	V	Y	F	L	L	D	S	R	
TRV-CAM- 55.	K	T	N	C	P	Y	S	A	F	K	R	F	P	D	K	S	L	Y	S	V	L	S	K	E	
TMV- 51.	V	W	K	P	S	P	Q	V	T	V	R	F	P	D	S	D	-	F	K	V	Y	R	Y	N	
SHMV- 51.	S	Q	V	S	V	V	S	P	I	S	R	F	P	A	E	P	A	Y	Y	I	Y	L	R	D	
BNYVV- 63.	D	S	N	P	F	V	S	P	M	T	R	F	P	Q	T	L	T	M	Y	G	A	L	V	L	
BSMV-166.	L	Q	S	A	R	L	P	V	Y	T	R	K	T	F	E	R	E	L	A	L	E	W	I	I	P
TRV-PSG-161.	L	E	N	M	Q	L	K	S	Y	Y	R	R	N	F	E	K	N	F	G	V	K	F	G	S	A
TRV-CAM-164.	L	E	G	L	S	L	R	P	T	Y	R	R	K	F	E	K	D	M	N	V	K	W	G	G	S
TMV-131.	E	L	I	R	G	T	G	S	Y	N	R	S	S	F	E	S	S	S	G	L	V	W	T	S	G
SHMV-132.	L	L	T	N	G	T	G	V	F	N	R	T	S	F	E	S	A	S	G	L	W	L	V	T	T
BNYVV-155.	A	L	R	L	E	E	L	M	W	T	R	D	K	F	E	D	R	F	K	L	P	W	T	P	V

Fig. 4. Comparison of the best-conserved amino acid sequences within the coat proteins of tobamoviruses (TMV, SHMV), TRV, BSMV, and BNYVV (Bouzoubaa et al. 1986; Cornelissen et al. 1986; Gustafson and Armour 1986; Altschuh et al. 1987). The number in front of a sequence block gives the position within the protein of the first amino acid of this block. Residues predicted to be involved in the intrasubunit salt bridge (see Altschuh et al. 1987) are indicated by asterisks. Positions where two or more sequences have conserved similarities as given by Schwartz and Dayhoff (1978) are indicated by boxes.

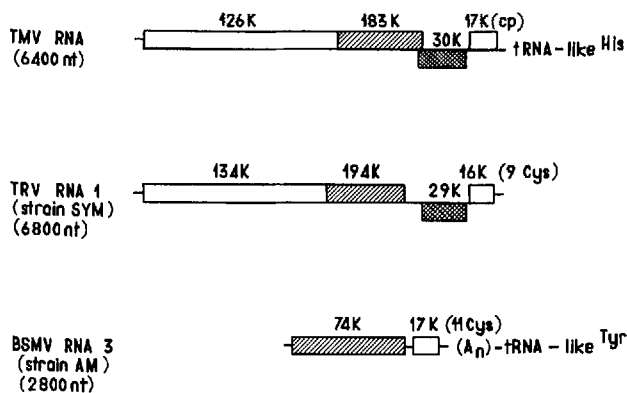


Fig. 5. Organization of TMV RNA (adapted from Palukaitis and Zaitlin 1986), TRV RNA 1 (from Cornelissen et al. 1986; Hamilton et al. 1987), and BSMV RNA 3 (Rupasov et al. 1986; Gustafson et al. 1987). Proteins containing similarities are marked by identical shading.

among the proteins of the capped supergroup RNA plant viruses was created in accordance with the maximum topological similarity principle of Yushmanov and Chumakov (1988). It is important to note that similarity between polymerase proteins encoded in BSMV RNA 3 and BNYVV RNA 1 is much less pronounced (Fig. 6) in spite of the clear relations in gene organization and the sequences of a number of other proteins between these viruses (see above).

The deduced phylogenetic tree revealed a significant divergence of the potexviruses from the other capped plant RNA viruses (Fig. 6B). Surprisingly, comparison of the PVX putative polymerase domain and the corresponding region of the 206-kd protein of turnip yellow mosaic virus (TYMV—a type member of spherical tymoviruses) whose RNA sequence was recently determined (Morch et al. 1988) leads to the conclusion that these viral polymerase proteins are quite closely related (Fig. 6). Thus, the gene coding for the putative RNA polymerase in the potexvirus and the tymovirus monopartite genomes retains phylogenetic relationships that are independent of adjacent genomic elements of the same RNA because PVX and TYMV capsid proteins are unrelated and the above-mentioned conserved triple block of genes is not found in the TYMV genome.

BSMV RNA 3: Cysteine-Rich Protein Gene

The 17-kd protein encoded in the 3'-terminal region of BSMV RNA 3 (Dolja et al. 1983a; Rupasov et al. 1986; Gustafson et al. 1987) resembles the 16-kd protein of TRV (Cornelissen et al. 1986) as well as the 14-kd protein of BNYVV (Bouzoubaa et al. 1986): all these proteins contain 6–8% cysteine residues dispersed along the polypeptide chain. The

hypothetical cysteine-rich proteins of the three viruses are encoded in the 3'-terminal regions of different genomic RNAs: TRV RNA 1, BNYVV RNA 2, and BSMV RNA 3. Comparative analyses of the cysteine-rich protein sequences reveal obvious relationships between these proteins and the proteins having metal-binding nucleic acid-binding domains (Fig. 7). The metal-binding “fingers” involve two cysteines separated by 5–25 residues from a brace of closely spaced histidines and/or cysteines (Klug and Rhodes 1987) (Fig. 7). However, the identification of the above-mentioned proteins as metal-binding proteins is rather questionable with respect to the recent experimentally based conclusion of Freedman et al. (1988): “Certainly the mere presence of appropriately placed cysteines or histidines should not suffice to identify zinc finger proteins.”

Between the 194-kd and 16-kd protein genes of TRV there is a gene for a 29-kd protein similar to the 30-kd transport protein of TMV (Boccaro et al. 1986) (see Fig. 5). Such entities are lacking in the BNYVV, WCIMV, PVX, and, possibly, BSMV genomes. In these cases the transport function can supposedly be performed by one or more of the above-mentioned triple block of proteins that are in the three viruses (see Fig. 2).

Discussion

To sum up, kindred genomic elements can be found in diverse and quite remote groups of plant viruses. Taking BSMV as a reference, for its tyrosine-accepting tRNA-like structure the related genomic elements are bromo- and cucumoviruses; for the conserved domains 1C and 2 of the putative RNA polymerases they are tobra- and tobamoviruses; for the triple gene block in RNA 2 they are furo- and potexviruses; for the coat-protein gene they are furo-, tobamo-, and tobaviruses. On the other hand, the PVX coat protein is akin to the potyvirus coat proteins. The last observation suggests that the different genes of the monopartite PVX genome are related to the genomic elements of dissimilar plant viruses, as is the case for the multipartite BSMV genome.

Recently, Miller et al. (1988) have sequenced the monopartite RNA genome of barley yellow dwarf virus (BYDV—a type member of the spherical luteoviruses). The BYDV genome unexpectedly was found to code for the 50-kd protein similar to the BNYVV 54-kd protein—the read-through product of the coat protein gene (see Fig. 2). The other protein of BYDV (60kd) displays considerable similarity with the putative RNA polymerase of carnation mottle virus (CarMV) (Miller et al. 1988). However, the genome strategy of BYDV and the sequences of

```

*** ***** * * *** ** * * * * *
BSMV 404. aLEIDFSKFDKSkTg1Hikavig1YKLFg...8...WEkSqyQTYVKDrnfGLeAyLLYQQKS
TRV 1451. fvEIDmSKFDKsANrFHLQlqLEIYRLFg...8...WEvSHtQTTyrDIqnGmmAHIWYQQKS
TMV 1382. vLEIDiSkYDKSQNeFhcavEyEIwRRLG...8...WkqgHRnTT1kDytaGikTcIWYQRKS
BNYVV 1837. ngvIDaaAcDsgQgvFtqliErhIYaaLG...7...WyfSfReKYVmqSer-yvrAHmSYvktS
PVX 1238. sLANDYTAfDQSQDgAmLOFEVLKAKhHc...4...IQaYIdIKTspQI--FLdT-LSIMRLT
WC1MV 1073. cfsNDfTAfDQSQDgSiLQFEVikAKfHN...5...IEGYIQIKThakI--FLGT-LSIMRLS
TYMV 1574. kiANDYTAfDQSQhGesvvLEALKmKRLN...5...IQLhvh1KInvst--qfGp-LtcMRLT

* *** * *** ***** * * *** ** * * *
BSMV GncDTYgsNTwsaA1ALldcLPLEdAhfcVFGGDDSLIIFDqGyiISDPCr--qLAgT-WNLE
TRV GDaDTYnANsDrtLcALLSeLPLEKAvmvtYGGDDSLIaFPrGtqFvDPCP--KLATk-WNFE
TMV GDvtTFiGNtViiAAcLaSMLPmEKiikgAfCgDDSLlyFPkGeeFpDvqh---SAnlmWNFE
BNYVV GEPGT1lGNtIlmgAmLnaMLagtgpfcmAmkGDDgfkqrqan-LKInDqML--KLikKetvLd
PVX GEGPTFDANTEcNIAYTHTKFDIPagtaQVYAGDDSaLDcvpevKhSFhrLEdKLILKsKpVf
WC1MV GEGPTFDANTEaNIAYTHTKFnIPedaAQVYAGDDmsIDyvasvKpSFnMiEhlmkLkGkPVf
TYMV GEPGTyDdNTDyNLAviySqyDvg-sepimvsGDDSLIDhP--LptrhdwP---SvLKr1hLr

** * ***** ***** * *****
BSMV CKvgDfkyP-aFCGKFLlc...6...FVPDaaKfITKLGrtD
TRV CKiFKydVP-mFCGKFLlk...6...FVPDPvKvLTKLGkKs
TMV aKLFKKqYg-yFCGrYVih...6...VyyDPLKLIsKLGAKh
BNYVV FKL-DLnVPitFCGYaLSn...2...lfPsvsrkLTKIaAhr
PVX tqQ-KKGSwPEFCGwLItP...2...VmKDPiKLhvSLkLae
WC1MV ntq-tqGdfaeFCGwtISP...2...iKkPeKmmSIElqk
TYMV FKL-eLtShPlFCGYyvGp...2...cIrnPLaLfcKLmiAv
A

```

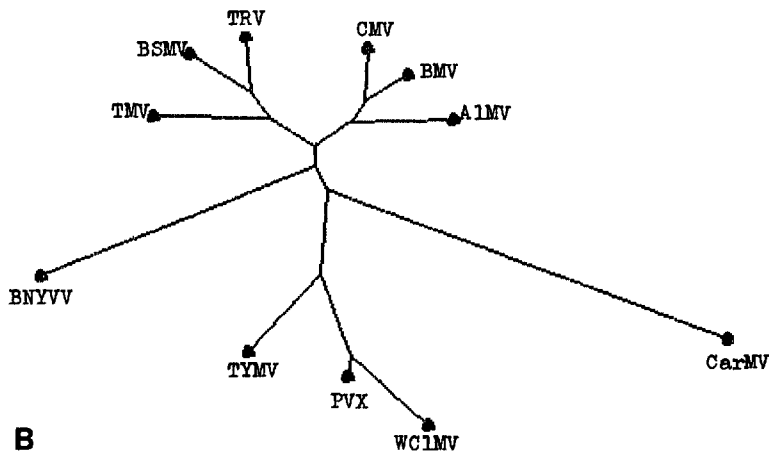


Fig. 6. **A** Comparison of the conserved regions of the putative RNA-dependent RNA polymerases of BSMV (Rupasov et al. 1986), TRV (Cornelissen et al. 1986), TMV (Goelet et al. 1982), BNYVV (Bouzoubaa et al. 1987), WC1MV (Forster et al. 1988), TYMV (Morch et al. 1988), and PVX. Identities in two or more sequences are indicated with capital letters. Asterisks denote identity between BSMV and TRV. Gaps (-) were inserted to facilitate comparison. Numerals show the numbers of amino acid residues that could not be aligned. The numbers in front of the sequences give the positions within the proteins of the first amino acid. The numbers of identical positions in the pairwise alignments are PVX vs WC1MV, 91; BSMV vs TRV, 76; TMV vs TRV, 64; BSMV vs TMV, 51; PVX vs TYMV, 51; BSMV vs BNYVV, 32; BSMV vs PVX, 26; TMV vs BNYVV, 33; TMV vs TYMV, 33; TMV vs PVX, 29; BNYVV vs TYMV, 30; BNYVV vs PVX, 25. **B** Maximum topological similarity rootless tree representing phylogenetic relationships among conserved domains of RNA polymerase-related sequences of capped plant (+)RNA-containing viruses (Chumakov and Yushmanov 1988). The sequences from TMV, BSMV, TRV, BNYVV, PVX, cucumber mosaic virus (CMV), BMV, AIMV, WC1MV, TYMV, and carnation mottle virus (CarMV) were aligned according to Koonin et al. (1987). Parameters of tree and matrices: mean square of relative distance deviation is 0.00200789; sum of distance deviation modes is 2.052; sum of differential neighbor matrix elements is 70; topologic mismatch is 5.45455%. Position of BNYVV in the tree fluctuates somewhat (local topological deviation is 16 for BNYVV and for other viruses it is 3-9).

other putative proteins differ significantly from those of the rodlike multipartite BNYVV and the spherical monopartite CarMV. Thus, BYDV represents one more virus whose monopartite genome consists

of a mosaic of genetic elements having independent phylogenetic relationships.

Much the same situation has been reported for animal viruses. Similarity has been found in the

TRV RNA 1	(1)	M	T	C	V	L	K	G	C	V	...	9aa...	T	C	S	I	G	H	A	N	...
TRV RNA 1	(45)	N	N	C	G	W	F	V	C	V	...	11aa...	C	C	G	R	S	H	L	E	...
BNYVV RNA 2	(66)	V	S	C	V	K	-	C	C	Q	...	9aa...	N	H	S	K	-	H	V	Q	...
BNYVV RNA 2	(82)	K	N	H	S	K	-	-	H	V	...	15aa...	V	C	G	D	C	C	E	S	...
BSMV RNA 3	(4)	F	S	C	V	-	-	C	C	G	...	10aa...	R	C	E	R	K	H	V	Y	...
BSMV RNA 3	(57)	H	S	C	G	M	-	L	C	S	...	14aa...	F	C	G	Q	K	H	A	D	...

Fig. 7. Amino acid sequence similarity within the cysteine-rich 16-kd protein of TRV (Cornelissen et al. 1986), the 17-kd protein of BSMV (Rupasov et al. 1986; Gustafson et al. 1987) and the 14-kd protein of BNYVV (Bouzoubaa et al. 1986). The conservative cysteine and histidine residues are boxed. Gaps (-) indicate nonconserved spacing between conserved positions. The numbers within and in front of the sequence blocks are defined in Fig. 1.

three-dimensional structure and in the amino acid sequences of the coat proteins of the members of picorna- and alphaviruses (Fuller and Argos 1987)—groups markedly differing in genome structure and mode of expression and belonging to different supergroups of RNA-containing viruses. Similar results were obtained upon comparison of the coat proteins of picornaviruses and small spherical (+)RNA-containing plant viruses (Rossmann 1987). Thus, every gene (or sometimes a block of genes) or noncoding regulatory genomic element (e.g., the tRNA-like structure) connotes its own scheme of evolutionary relations. This complicates elaboration of a universal system of (+)RNA-containing viruses reflecting their actual evolutionary links. Such a system can be envisaged no longer as a tree but rather as a more intricate structure, a kind of three-dimensional network.

It is only natural to suggest that such involved patterns may result from real processes of exchange and reassortment of genetic information among viruses. Recent studies also facilitate perception of the mechanism of this reassortment: it is recombination at the RNA level, most likely through copy-choice during replication. Such recombination has been determined for at least three groups of (+)RNA-containing viruses: picorna- (Kirkegaard and Baltimore 1986; Romanova et al. 1986), corona- (Makino et al. 1986), and bromoviruses (Bujarsky and Kaesberg 1986). Especially demonstrative is the last example, describing an exchange of tRNA-like structures between different genomic RNAs or BMV. It is tempting to think that recombination under natural conditions also gave rise to the RNA of SHMV whose coding body carries the same genes as in other tobamoviruses, whereas the tRNA-like structure is similar to that of tymoviruses (Meshi et al. 1981).

Findings of duplications and a deletion in the putative BSMV RNA polymerase gene (Dolja et al. 1983b; Afanasiev et al. 1986; Gustafson et al. 1987) indirectly point to the existence of a process involving copy-choice in this virus as well. Such events may produce new BSMV strains both in nature and under experimental conditions (Atabekov and Dolja 1986). The peculiar 3'-terminal tandem in BSMV

RNAs 1–3 could itself have arisen from recombination between an RNA carrying a 3'-terminal tRNA-like structure (as in bromoviruses) and an RNA carrying a 3'-terminal poly(A) tail (as in PVX or BNYVV).

It can be supposed that the genome of RNA-containing viruses is composed of discrete units that may differ in size from several dozen (noncoding regulatory sequences) to several thousand nucleotides (blocks of genes). The unit size may itself be determined by its function. At the same time, the location and environment of each unit may vary greatly.

Assuming that formation of viral RNA genomes involved reassortment of their elements, we come to the next question of what host cells exist wherein these genomes could meet. At present the answer can be speculative only. The two versions that seem the most plausible are: (1) viruses and host diverged concurrently, i.e., the common ancestors of their contemporary viruses infected the common ancestors of their contemporary hosts; (2) the host specificity of viruses and/or the specific susceptibility of their hosts is a later development than was the formation (establishment) of the viral RNA genome.

There appear to be some uncertainties in the interpretation of these phenomena. As suggested by Matthews (1985): (1) "... the genomes of viruses with RNA involved in part or all of the genomic life cycle may consist of a mixture of highly conserved genes and genes with no base sequence homology due to the rapid rate of mutation in RNA. . . ."; (2) "Quite unrelated viruses that evolved independently might have developed amino acid homologies in those parts of the protein constrained to interact with some highly conserved host macromolecule or viral process for which there is only one structural solution."

There is another plausible cause of similarity among genomic elements of distantly related viruses, which is capture of similar sequences from the host cell. Although for coding viral sequences the examples are still scarce (except for retroviruses) and not always convincing, alleged acquisition of a poly(A) tail or a tRNA-like structure in just this

manner could hardly be disputed. Entrapment and covalent attachment of cell tRNA has been described for Sindbis virus DI RNA (Monroe and Schlesinger 1983). It is noteworthy that both animal and plant cell genomes (Sakamoto and Okada 1985; Benslimane et al. 1986) contain numerous transcribed sequences encoding not only tRNAs but also tRNA-like structures. The exchange of information between cell and viral RNAs through a copy-choice mechanism may be facilitated by the presence of endogenous RNA-dependent RNA polymerases in animal (Volloch 1986) and plant cells (Fraenkel-Conrat 1983).

By and large, analysis of genome structure of a number of elongated (+)RNA-containing plant viruses strongly supports the suggestion that shuffling of genomic elements among different viruses has occurred. As far as we know, such an idea was first mentioned by Haseloff et al. (1984).

There are a number of convincing examples in favor of reassortment between the genomes of RNA-containing viruses differing in particle structure and mode of genome replication and expression. Recent findings that the products of cauliflower mosaic virus genes I and II are somewhat related to the proteins of several RNA-containing plant viruses (Hull et al. 1986; Domier et al. 1987) bring the DNA-containing reovirus into the field of play as well. At least one example of amino acid sequence similarity exists between the coat proteins of single-stranded DNA-containing and RNA-containing viruses (Mohana Rao 1986). Thus, the overall impression is that among viruses the shuffling of genetic elements is rather a universal phenomenon.

Acknowledgments. We are grateful to Dr. K.M. Chumakov for computer programs and to Drs. K.M. Chumakov and E.V. Koonin for helpful discussions and for providing the results of their investigations prior to publication. Thanks are due to Dr. A.V. Galkin for critically reading the manuscript.

References

- Afanasiev BN, Rupasov VV, Fedchenko VM, Golova YS, Dolja VV, Atabekov JG, Kozlov YV, Bayev AA (1986) The primary structure of barley stripe mosaic virus RNA 3 and formation of RNA 2b and RNA 4. *Dokl Akad Nauk SSSR* 291: 742-746
- Agranovsky AA, Dolja VV, Gorbulev VG, Kozlov YV, Atabekov JG (1981) Aminoacylation of barley stripe mosaic virus RNA: polyadenylate-containing RNA has a 3' terminal tyrosine-accepting structure. *Virology* 113:174-187
- Agranovsky AA, Dolja VV, Atabekov JG (1982) Structure of the 3' extremity of barley stripe mosaic virus RNA: evidence for internal poly(A) and a 3' terminal tRNA-like structure. *Virology* 119:51-58
- Ahluquist P, Luckow V, Kaesberg P (1981) Complete nucleotide sequence of brome mosaic virus RNA 3. *J Mol Biol* 153: 23-38
- Allison R, Johnston RE, Dougherty WG (1986) The nucleotide sequence of the coding region of tobacco etch virus genomic RNA: evidence for the synthesis of a single polypeptide. *Virology* 154:9-20
- Altschuh D, Lesk AM, Bloomer AC, Klug A (1987) Correlation of co-ordinated amino acid substitutions with function in viruses related to tobacco mosaic virus. *J Mol Biol* 193:693-707
- Argos P, Kamer G, Nicklin MJH, Wimmer E (1984) Similarity in gene organization between proteins of animal picornaviruses and plant comoviruses suggest common ancestry of these virus families. *Nucleic Acids Res* 12:7251-7267
- Atabekov JG, Dolja VV (1986) Hordeiviruses: structure and replication. In: Van Regenmortel MHV, Fraenkel-Conrat H (eds) *The plant viruses*, vol 2. Plenum, New York, pp 397-420
- Benslimane AA, Dron M, Hartman C, Rode A (1986) Small tandemly repeated DNA sequences of higher plants likely originate from a tRNA gene ancestor. *Nucleic Acids Res* 14: 8111-8119
- Boccaro M, Hamilton WDO, Baulcombe DC (1986) The organization and intervirial homologies of genes at the 3' end of tobacco rattle virus RNA 1. *EMBO J* 5:223-229
- Bouzoubaa S, Ziegler V, Beck D, Guilley H, Richards K, Jonard G (1986) Nucleotide sequence of beet necrotic yellow vein virus RNA-2. *J Gen Virol* 67:1689-1700
- Bouzoubaa S, Quillet L, Guilley H, Jonard G, Richards K (1987) Nucleotide sequence of beet necrotic yellow vein virus RNA-1. *J Gen Virol* 68:615-626
- Bujarsky JE, Kaesberg P (1986) Genetic recombination between RNA components of a multipartite plant virus. *Nature* 231: 528-531
- Chumakov KM, Yushmanov SV (1988) Maximum topological similarity principle in molecular taxonomy. *Mol Genetika* 3: 3-9
- Cornelissen BJC, Bol JF (1984) Homology between the proteins encoded by tobacco mosaic virus and two tricornaviruses. *Plant Mol Biol* 3:379-384
- Cornelissen BJC, Linthorst HJM, Brederode F, Bol JF (1986) Analysis of the genome of tobacco rattle virus strain PSG. *Nucleic Acids Res* 14:2157-2169
- Dever TE, Glynias MJ, Merrick WC (1987) GTP-binding domains: three consensus sequence elements with distinct spacing. *Proc Natl Acad Sci USA* 84:1814-1818
- Dolja VV, Agranovsky AA, Lunina NA, Atabekov JG (1983a) Short virion RNA in barley stripe mosaic virus. *FEBS Lett* 151:215-218
- Dolja VV, Lunina NA, Leiser R-M, Stanarius T, Belzhelarskaya SN, Kozlov YV, Atabekov JG (1983b) A comparative study on the in vitro translation products of individual RNAs of two-, three-, and four-component strains of barley stripe mosaic virus. *Virology* 127:1-14
- Dolja VV, Grama DP, Morozov SY, Atabekov JG (1987) Potato virus X-related single- and double-stranded RNAs. *FEBS Lett* 214:308-312
- Domier LL, Shaw JG, Rhoads RE (1987) Potyviral proteins share amino acid sequence homology with picorna-, como-, and caulimoviral proteins. *Virology* 158:20-27
- Eisenberg D, Schwarz E, Komoromy M, Wall R (1984) Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J Mol Biol* 179:125-142
- Forster RLS, Bevan MW, Harbison S-N, Gardner RC (1988) The complete nucleotide sequence of the potyvirus white clover mosaic virus. *Nucleic Acids Res* 16:291-303
- Fraenkel-Conrat H (1983) RNA-dependent RNA-polymerases of plants. *Proc Natl Acad Sci USA* 80:422-424
- Franssen H, Leunissen J, Goldbach R, Lomonosoff GF, Zimmern D (1984) Homologous sequences in non-structural proteins

- from cowpea mosaic virus and picornaviruses. *EMBO J* 3: 855-861
- Freedman LP, Yamamoto KR, Luisi BF, Sigler PB (1988) More fingers in hand. *Cell* 54:445-446
- Fuller SD, Argos P (1987) Is Sindbis a simple picornavirus with an envelope? *EMBO J* 6:1099-1105
- Goelet P, Lomonosoff GP, Butler PJG, Akam ME, Gait MJ, Karn J (1982) Nucleotide sequence of tobacco mosaic virus RNA. *Proc Natl Acad Sci USA* 79:5818-5822
- Goldbach RW (1986) Molecular evolution of plant RNA viruses. *Annu Rev Phytopathol* 24:289-310
- Gorbalenya AE, Blinov VM, Koonin EV (1985) Prediction of nucleotide-binding properties of virus-specific proteins from their primary structure. *Mol Genetika* 11:30-36
- Gorbalenya AE, Blinov VM, Donchenko AP, Koonin EV (1988a) An NTP-binding motif is the most conserved sequence in a highly diverged monophyletic group of proteins involved in positive strand RNA viral replication. *J Mol Evol* (in press)
- Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1988b) A conserved NTP-motif in putative helicases. *Nature* 333:22
- Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1988c) A novel superfamily of nucleoside triphosphate-binding motif containing proteins which are probably involved in duplex unwinding in DNA and RNA replication and recombination. *FEBS Lett* 235:16-24
- Guilford PJ, Forster RLS (1986) Detection of polyadenylated subgenomic RNAs in leaves infected with potexvirus daphne virus X. *J Gen Virol* 67:83-90
- Gustafson GD, Armour SL (1986) The complete nucleotide sequence of RNA-beta from the type strain of barley stripe mosaic virus. *Nucleic Acids Res* 14:3895-3909
- Gustafson GD, Hunter B, Hanau R, Armour SL, Jackson AO (1987) Nucleotide sequence and genetic organization of barley stripe mosaic virus RNA gamma. *Virology* 158:394-406
- Haenni AL, Joshi S, Chapeville F (1982) Transfer RNA-like structures in the genomes of RNA viruses. *Prog Nucleic Acid Res Mol Biol* 27:85-104
- Hamilton WDO, Boccara M, Robinson DJ, Baulcombe DC (1987) The complete nucleotide sequence of tobacco rattle virus RNA-1. *J Gen Virol* 68:2563-2575
- Harrison BD, Robinson DJ (1986) Tobraviruses. In: Van Regenmortel MHV, Fraenkel-Conrat H (eds) *The plant viruses*, vol 2. Plenum, New York, pp 339-372
- Haseloff J, Goelet P, Zimmern S, Ahlquist P, Dasgupta R, Kaesberg P (1984) Striking similarities in amino acid sequence among nonstructural proteins encoded by RNA viruses that have dissimilar genomic organization. *Proc Natl Acad Sci USA* 81:4358-4362
- Hodgman TC (1988) A new superfamily of replicative proteins. *Nature* 333:22-23
- Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci USA* 78:3824-3828
- Hull R, Sadler J, Longstaff M (1986) The sequence of carnation etched ring virus DNA: comparison with cauliflower mosaic virus and retroviruses. *EMBO J* 5:3083-3090
- Kamer G, Argos P (1984) Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res* 12:7269-7282
- Kirkegaard K, Baltimore D (1986) The mechanism of RNA recombination in poliovirus. *Cell* 47:433-443
- Klug A, Rhodes D (1987) "Zinc fingers": a novel protein motif for nucleic acid recognition. *Trans Int Biochem Soc* 12:464-469
- Koonin EV, Gorbalenya AE, Chumakov KM, Donchenko AP, Blinov VM (1987) Evolution of RNA-dependent RNA-polymerases of positive strand RNA viruses. *Mol Genetika* 7:27-39
- Koonin EV, Chumakov KM, Yushmanov SV, Gorbalenya AE (1988) Evolution of RNA-dependent RNA-polymerases of positive strand RNA viruses: a comparison of phylogenetic trees generated by different methods. *Mol Genetika* 3:16-19
- Kozlov YV, Rupasov VV, Adyshev VD, Belgelarskaya SN, Agranovsky AA, Mankin AS, Morozov SY, Dolja VV, Atabekov JG (1984) Nucleotide sequence of the 3' terminal tRNA-like structure in barley stripe mosaic virus genome. *Nucleic Acids Res* 12:4001-4009
- Krayev AS, Morozov SY, Lukasheva LI, Rosanov MN, Chernov BK, Simonova ML, Golova YB, Belzhelarskaya SN, Pozmogova GE, Skryabin KG, Atabekov JG (1988) The complete nucleotide sequence and organization of potato virus X genome. *Dokl Akad Nauk SSSR* 300:711-716
- Lemaire O, Merdingoglu D, Valentin P, Putz C, Ziegler-Graff F, Guillely H, Jonard G, Richards K (1988) Effect of beet necrotic yellow vein virus RNA composition on transmission by *Polymyxa betae*. *Virology* 162:232-235
- Makino S, Keck JG, Stohlman SA, Lai MM (1986) High-frequency RNA recombination of murine coronaviruses. *J Virol* 57:729-737
- Matthews REF (1985) Viral taxonomy for the nonvirologist. *Annu Rev Microbiol* 39:451-474
- Meshi T, Ohno T, Iba H, Okada Y (1981) Nucleotide sequence of a cloned cDNA copy of TMV (cowpea strain) RNA, including the assembly origin, the coat protein cistron, and the 3' non-coding region. *Mol Gen Genet* 184:20-25
- Meyer M, Hemmer O, Mayo MA, Fritsch C (1986) The nucleotide sequence of tomato black ring virus RNA-2. *Gen Virol* 67:1257-1271
- Miller WA, Waterhouse PM, Gerlach WL (1988) Sequence and organization of barley yellow dwarf virus genomic RNA. *Nucleic Acids Res* 16:6097-6112
- Mohana Rao JK (1986) Sequence homology among the coat proteins of geminiviruses. *Biochem Biophys Res Commun* 134:372-377
- Mohana Rao JK, Argos P (1986) A conformational preference parameter to predict helices in integral membrane proteins. *Biochim Biophys Acta* 869:197-214
- Monroe SS, Schlesinger S (1983) RNAs from two independently isolated defective-interfering particles of Sindbis virus contain a cellular tRNA sequence at their 5' ends. *Proc Natl Acad Sci USA* 80:3279-3282
- Morch M-D, Boyer J-C, Haenni A-L (1988) Overlapping open reading frames revealed by complete nucleotide sequencing of turnip yellow mosaic virus genomic RNA. *Nucleic Acids Res* 16:6157-6190
- Morozov SY, Gorbulev VG, Novikov VK, Agranovsky AA, Kozlov YV, Atabekov JG, Bayev AA (1981) The primary structure of the 5'- and 3'-terminal regions of the genomic RNA of potato virus X. *Dokl Akad Nauk SSSR* 259:723-725
- Morozov SY, Lukasheva LI, Chernov BK, Skryabin KG, Atabekov JG (1987) Nucleotide sequence of the open reading frames adjacent to the coat protein cistron in potato virus X genome. *FEBS Lett* 213:438-442
- Palukaitis P, Zaitlin M (1986) Tobacco mosaic virus: infectivity and replication. In: Van Regenmortel MHV, Fraenkel-Conrat H (eds) *The plant viruses*, vol 2. Plenum, New York, pp 105-132
- Quadt R, Verbeek HJM, Jaspars EMJ (1988) Involvement of a nonstructural protein in the RNA synthesis of brome mosaic virus. *Virology* 165:256-261
- Romanova LI, Blinov VM, Tolskaya EA, Victorova EG, Kolesnikova MS, Guseva EA, Agol VI (1986) The primary structure of crossover regions of intertypic poliovirus recombi-

- nants: a model of recombination between RNA genomes. *Virology* 155:202-213
- Rossmann MG (1987) Evolution of molecular structure in spherical RNA viruses. *BioEssays* 7:99-103
- Rupasov VV, Afanasiev BN, Adyshev DM, Kozlov YV (1986) The primary structure of the 3'-terminal regions of RNA 1 and RNA 3 of barley stripe mosaic virus. *Dokl Akad Nauk SSSR* 288:1237-1241
- Sakamoto K, Okada N (1985) Rodent type 2 Alu family, rat identifier sequence, rabbit c family and bovine or goat 73-bp repeat have evolved from tRNA genes. *J Mol Evol* 22:139-140
- Schwartz RM, Dayhoff MO (1978) In: Dayhoff MO (ed) *Atlas of protein sequence and structure*, vol 5, suppl 3. National Biomedical Research Foundation, Washington DC, pp 353-358
- Sonenberg N, Shatkin AJ, Riccardi PP, Rubin M, Goodman RM (1978) Analysis of terminal structures of RNA from potato virus X. *Nucleic Acids Res* 5:2501-2512
- Staden R (1982) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. *Nucleic Acids Res* 10:2951-2961
- Symons RH (1985) Viral genome structure. In: Van Regenmortel MHV, Fraenkel-Conrat H (eds) *The plant viruses*, vol 1. Plenum, New York, pp 57-82
- Volloch V (1986) Cytoplasmic synthesis of globin mRNA in differentiated murine erythroleukemia cells: possible involvement of RNA-dependent RNA polymerase. *Proc Natl Acad Sci USA* 83:1208-1212
- Yushmanov SV, Chumakov KM (1988) Algorithms for construction of maximum topological similarity phylogenetic trees. *Mol Genetika* 3:9-15

Received July 8, 1988/Revised October 26, 1988

Note added in proof. The demonstration of nonstringent specificity of interaction between tobamovirus replicase component(s) and the 3' noncoding region of viral RNA [Ishikawa M, Meshi T, Watanabe Y, Okada Y (1988) Replication of chimeric tobacco mosaic viruses which carry heterologous combinations of replicase genes and 3' noncoding regions. *Virology* 164:290-293] allows us to suggest that the process of genome recombination between riboviruses may not be so uncommon. Recently, the strong indications on the recombination between distantly related or unrelated viral RNA genomes have been revealed under natural conditions for animal coronaviruses and orthomyxovirus [Luytjes W, Bredenbeek PJ, Noten AFH, Horzinek MC, Spaan WJM (1988) Sequence of mouse hepatitis virus A59 mRNA: indications for RNA recombination between coronaviruses and influenza C virus. *Virology* 166:415-422] and for plant luteovirus and sobemovirus [Veidt, I, Lot H, Leiser M, Scheidecker D, Guilley H, Richards K, Jonard G (1988) Nucleotide sequence of beet western yellows virus RNA. *Nucleic Acids Res* 16:9917-9932]. Possible evolutionary mechanisms involving gene reassortment between viral RNAs have also been discussed in the recent reviews [Gibbs A (1987) Molecular evolution of viruses: "trees," "clocks" and "modules." *J Cell Sci Suppl* 7:319-337]; Zimmern D (1988) Evolution of RNA viruses. In: Holland J, Domingo E, Ahlquist P (eds) *RNA genetics*. CRC Press, Boca Raton, FL; and Goldenbach R, Wellink J (1988) Evolution of plus-strand RNA viruses. *Intervirology* 29: 260-267].