Alpha-like viruses in plants

R. Goldbach, Olivier O. Le Gall, J. Wellink

seminars in **VIROLOGY**, Vol 2, 1991 : pp 19-25

Alpha-like viruses in plants

Rob Goldbach, Olivier Le Gall and Joan Wellink**

Plant RNA viruses belonging to the alpha-like supergroup differ greatly in genome structure, translation strategy and capsid morphology. These viruses have much in common, however, in that their genomic RNAs possess a 5'-cap structure, that they produce subgenomic mRNAs and that they specify proteins exhibiting significant sequence homology to two non-structural proteins of Sindbis virus, containing a nucleotide binding and a polymerase domain, respectively. In addition, the plant viruses with a small RNA genome form two new distinct supergroups, the carmo-like supergroup that contains viruses that only share a conserved (carmovirus-related) polymerase domain, and the sobemo-like supergroup, encompassing viruses with conserved (sobemo-like) polymerase and putative protease domains.

Key words: RNA viruses / sequence homology / polymerase / supergroup / α-like / carmo-like / sobemo-like

VIRUSES with RNA genomes show a wide variation in particle morphology and genome structure and can parasitize prokaryotes (bacteriophages) as well as eukaryotes, both plants and animals. On the basis of their genome form they have been divided into the positive-stranded, the negative-stranded, and the double-stranded RNA viruses. The positive-stranded RNA viruses represent the largest and by far the most diverse group. Based on differences in genome organization, capsid morphology and biological properties such as host range, disease symptoms and epidemiology, the positive-stranded RNA viruses have been further subdivided into a large number of families (phages, animal viruses) or groups (plant viruses). Despite all these differences, nucleotide sequence data that have become available for many viruses to date, reveal that most of the eukaryotic positive-stranded RNA viruses, irrespective of whether they infect plants or animals, can be clustered into a limited number of 'supergroups'.[1-3] As previously described,[2,4] a large number of plant- and animal-infecting RNA viruses can thus be placed into two supergroups, the picorna-like viruses (all related to the Picornaviridae) and the Sindbis-like or α-like viruses (all related to the alphaviruses, a genus within the family Togaviridae). With more viral sequences having been published over the past two years the concept of 'supergrouping' has gained increasing support, although more than two supergroups now have to be considered. A third supergroup, centered around the plant carmo-, tombus- and luteoviruses ('carmo-like' supergroup) has been proposed as a first split-off from the supergroup of α-like viruses (refs 2,5,6; Table 1), and recently published sequence alignments indicate that the polymerase of the plant viruses belonging to this supergroup is more related to those of the Flaviviridae.[7,8] As discussed in this review a second split-off of the original α-like supergroup has to be envisaged, giving rise to an additional supergroup, the 'sobemo-like' viruses (Table 1). This review will mainly focus on the α-like plant viruses but will also discuss the evolutionary position of the small RNA-containing viruses of plants (the carmo-like and sobemo-like viruses), which appear to be less clearly related to the Togaviridae. The supergroup of picorna-like viruses is discussed in the article by King *et al*, in this issue.

The α-like plant viruses

As summarized in Table 1 the plant viruses belonging to the supergroup of α-like viruses, though widely variable in virion-structure and hosts, share a number of properties with the alphaviruses indicating a close genetic relationship. They all encode a number of non-structural proteins, specified by a similarly ordered gene set and they all have capped RNA genomes. Furthermore, although their translation strategies can differ widely, they all produce subgenomic mRNAs and their genomes often contain leaky termination codons, allowing the production of read-through proteins. In Figure 1 the genomes of a number of α-like plant viruses are compared with that of the best studied member of the alphaviruses, Sindbis virus.

*From the Departments of Virology and *Molecular Biology, Agricultural University, Binnenhaven 11, 6709 PD Wageningen, The Netherlands*

1044-5773/91/0201-0003$5.00/0

Table 1. Supergroups of positive-strand RNA viruses infecting plants

Supergroup	Group	RNA termini 5′	RNA termini 3′	Subgenomic mRNA	Conserved domains mtr hel pro pol*	Remarks
Picorna-like	Como Nepo Poty	VPg	Poly A	–	– + + +	Polyprotein processing
Alpha-like	Alfalfa mosaic Bromo Carla Clostero A Cucumo Furo Hordei Ilar Potex Tobamo Tobra Tymo	Cap	Poly A or tRNA or X_{OH}	+	+ + – +	Often readthrough
Carmo-like	Carmo Diantho Luteo (BYDV) MCMV Tobacco necrosis Tombus	Cap or VPg or ppX	X_{OH}	+	– – – +	Spherical small genome r/t or f/s*
Sobemo-like	Luteo Sobemo	VPg	X_{OH}	+	– – + +	Spherical small genome

*Abbreviations: hel, putative helicase; pro, (putative) proteinase; pol, (putative) polymerase; mtr, (putative) methyl transferase; r/t, readthrough; f/s, frame shifting.

From the comparison it is obvious that members of this supergroup are variations on one theme, all of them being very similar in genome organization and replication strategy. Although their genomes may be tripartite (e.g. BMV), bipartite (e.g. TRV) or undivided (e.g. TMV and Sindbis virus) they all specify proteins exhibiting significant sequence homology, encoded by a similarly arranged gene set (Figure 1).

These conserved proteins all seem directly involved in the RNA replication process. Although direct biochemical evidence for this involvement is still lacking, both genetic[9-12] and sequence alignment data[2,13] support this view. Hence, the gene set conserved among all α-like viruses may be regarded as an RNA replication gene module.

Firstly, one of the three conserved proteins (or protein domains) contains sequence motifs, especially a GDD motif (indicated ■ in Figure 1), characteristic of all RNA-dependent RNA polymerases.[1,13] Hence, it is generally believed that the conserved protein domains containing these motifs represent core RNA-dependent RNA polymerases. Strikingly, in some alphaviruses (e.g. Sindbis virus) and several plant (e.g. tobamo- and tobraviruses) this polymerase domain is translated upon ribosomal read-through at a suppressible amber termination codon, which further underlines the genetic relationship between these viruses.

A second protein (domain) conserved among all members of the α-like supergroup contains a nucleotide-triphosphate binding (NTB) motif (★ in Figure 1).[14] Further sequence comparisons, moreover, revealed that the NTB-motif-containing proteins of the α-like supergroup have distant homology to a family of *Escherichia coli* helicases, i.e. UvrD, Rep, RecB and RecD.[15-17] It has therefore been postulated that all viruses of the α-like supergroup specify proteins with helicase activity, a function that may be involved in the unwinding of replicative-form (RF) molecules during replication.[16]

In addition to a core polymerase and putative helicase all α-like virus share a third conserved protein (or protein domain), always 5′-terminally encoded in the genome (Figure 1). For Sindbis virus this conserved domain is located in nsP1 and genetic evidence suggests that this protein represents a methyltransferase, possibly involved in capping of

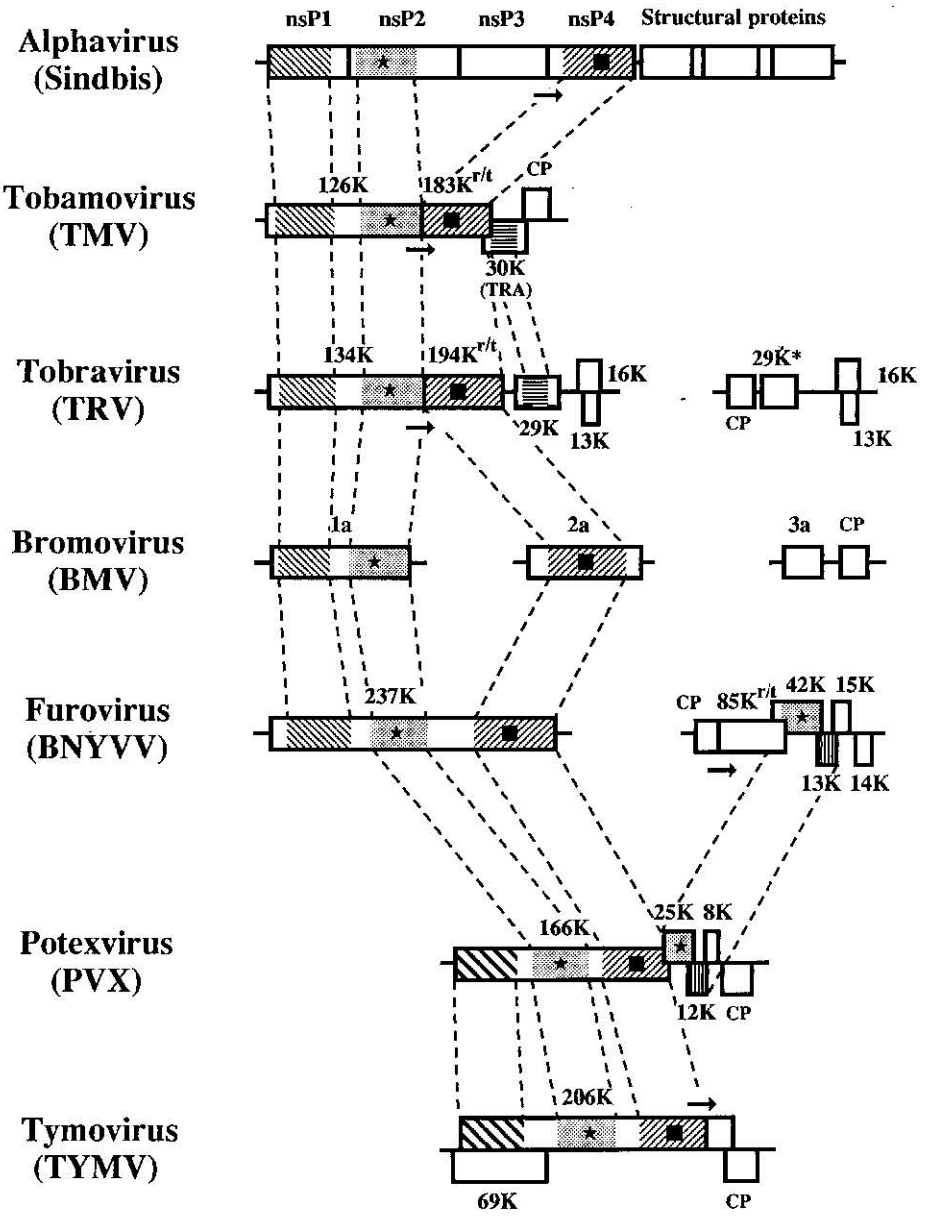

Figure 1. Comparison of the genomes of Sindbis virus (genus Alphavirus, family Togaviridae) and some α-like plant viruses. TMV, tobacco mosaic virus; TRV, tobacco rattle virus; BMV, brome mosaic virus; BNYVV, beet necrotic yellow vein virus; PVX, potato virus X; TYMV, turnip yellow mosaic virus. Coding regions in the genomes are indicated as open bars; regions of amino acid sequence homology in the gene products are indicated by similar shading. Other notations: CP, coat protein; TRA, transport protein; ★, nucleotide binding sequence motif; ■, conserved polymerase domain; →, leaky termination codon; r/t, readthrough protein.

the viral genome.[18] It would make sense for viruses with capped genomic RNAs (like all α-like viruses) to encode a specific capping enzyme since the cellular enzymes involved in the capping of mRNAs are primarily located in the nucleus and thus not available for viral RNA molecules replicating in the cytoplasm. Two subgroups can be distinguished with regard to this conserved 5′ domain within the α-like supergroup (Figure 1). There is clear interviral homology within, but hardly any between,

each of these subgroups.[19] They are respectively composed of α-, tobamo-, bromo-, cucumo-, hordei-, furo- and A1MV-groups on the one hand, and of potex-, tymo-, carla-, and closterovirus (type A) on the other. Interestingly enough, the same splitting of the α-like supergroups has recently been proposed,[20,21] based on core polymerase and NTB domains multiple sequence comparisons (see also below). However, whereas it can reasonably be assumed that this domain in the first subgroup, by analogy with Sindbis virus, codes for a methyltransferase, the same assessment for the other subgroup is more hazardous because of the lack of homology with Sindbis virus.

In summary, all three protein domains conserved among the α-like viruses are likely to be involved in the replication of the viral genome. In addition to these clearly conserved domains, α-like viruses also encode more specific proteins, both structural and non-structural.

With respect to the structural proteins it is not surprising that these are variable since this variation reflects adaptation of viruses to different environments and therefore different selective pressures; for example, the presence (vertebrates) or absence (lower animals, plants) of a host immune system. Unique non-structural proteins include nsP3 of Sindbis virus and the transport proteins (TRA, required for cell-to-cell spread of the virus; ref 22) encoded by plant virus members (e.g. TMV, BMV, TRV) of the supergroup. Furthermore, some viruses possess extra gene sets. For instance, the genomes of hordeivirus BSMV, furovirus BNYVV, potexvirus PVX and carlavirus PVS have acquired an additional gene set of which two genes are conserved (Figure 1). One of these conserved genes encodes a second protein with an NTB-motif (★ in Figure 1), while the coding product of the second conserved gene is not only conserved in terms of amino acid sequence but also in size (12-14 kDa).

The linkage of unique genes to a set of conserved genes indicates that the major differences between the members of the α-like supergroup are based on recombination events. As discussed in more detail elsewhere[4] the sequence information available to date indicates that recombination events have occurred within a single plant virus (e.g. BMV, TRV), between viruses belonging to the same taxonomic group (e.g. between tobraviruses TRV and PEBV) and between viruses belonging to different taxonomic groups (e.g. between luteo-, carmo- and sobemoviruses).

Re-grouping of the plant viruses with a small RNA genome

Although a number of plant viruses with a small RNA genome (a.o. carmo-, tombus-, and luteoviruses) have previously been placed into the supergroup of α-like viruses,[2] their evolutionary position needs to be reconsidered. With sequences of more of these small RNA viruses now being available it seems justified to distinguish these viruses from the 'true' α-like viruses on the basis of the following criteria (see also Table 1):

(1) the genomes of these viruses do not contain genes for a putative helicase (nor for a putative methyltransferase) (Figure 2).
(2) their polymerases are distinct, showing little or no sequence homology (outside the GDD motif and surrounding sequences) with the polymerases of the true alphavirus-related plant viruses.

Even within the conserved polymerase motifs the small RNA viruses form a cluster that is distinct from the true α-like viruses (and picorna-like viruses), as shown in the dendrogram of Figure 3 (taken from ref 21). Moreover, this hierarchical clustering of plant viral polymerases indicates that the group formed by carmo-, luteo- (BYDV), tombus and dianthoviruses on the one hand and the sobemo- and some other luteoviruses (PLRV, BWYV) on the other have distinct polymerases (Figure 3). It has previously been suggested to cluster the former group of small RNA viruses into a third supergroup, the carmo-like viruses.[2,5,6] Indeed it has been shown recently[7,8] that the polymerases of these plant viruses show some homology to the polymerase of flavi- and pestiviruses and hepatitis C virus, all belonging to the family of Flaviviridae, and clearly less homology to the polymerase of alphaviruses belonging to the family of Togaviridae.

The sobemoviruses are quite distinct from the carmo-like viruses, however, and certainly do not belong to this supergroup. First, the sobemoviral polymerase is not related to those of the carmo-like viruses (and neither to those of the α-like viruses (Figures 2 and 3). Furthermore, sobemoviruses have a 5′-VPg and specify a putative, trypsin-like serine proteinase (Figure 2; refs 23,24). Hence, these viruses seem to form a separate supergroup that may be intermediate to the picorna-like and carmo-like viruses. Surprisingly some luteoviruses specify a sobemo-like polymerase and protease (e.g. PLRV and BWYV), while another luteovirus (BYDV)

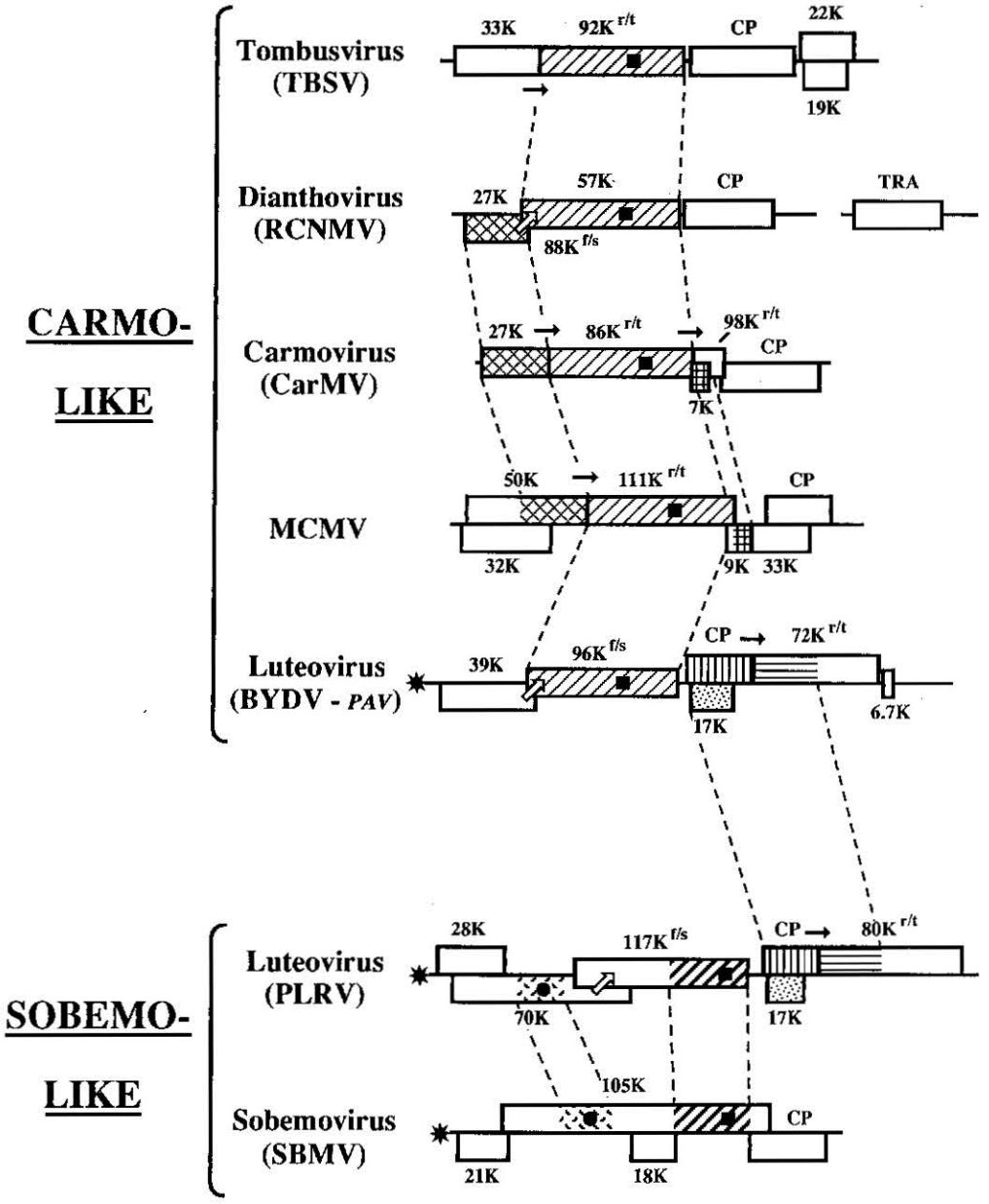

Figure 2. Comparison of the genomes of small, spherical RNA viruses of plants. TBSV, tomato bushy stunt virus; RCNMV, red clover necrotic mottle virus; CarMV, carnation mottle virus; MCMV, maize chlorotic mottle virus (unclassified); BYDV, barley yellow dwarf virus; PLRV, potato leafroll virus; SBMV, southern bean mosaic virus. Although the coat proteins of all these viruses possess regions of amino acid sequence homology, these regions are not shaded for reasons of simplicity. For symbols see Figure 1. Other notations: ⇨, ribosomal frame-shift; f/s, frame-shift protein; ✸, VPg, ●, putative proteinase domain.

specifies a very distinct polymerase homologous to that of the carmo-like viruses (Figures 2 and 3; refs 4,25,26). Since luteoviruses possess, like the sobemoviruses, a VPg-containing genome it is tempting to assume that the prototype luteoviral genome encodes a sobemo-like polymerase and that the odd genome of BYDV represents a recombinant RNA that has recently captured a carmo-like polymerase gene by interviral recombination. This idea is further strengthened by the recent finding, reported at the

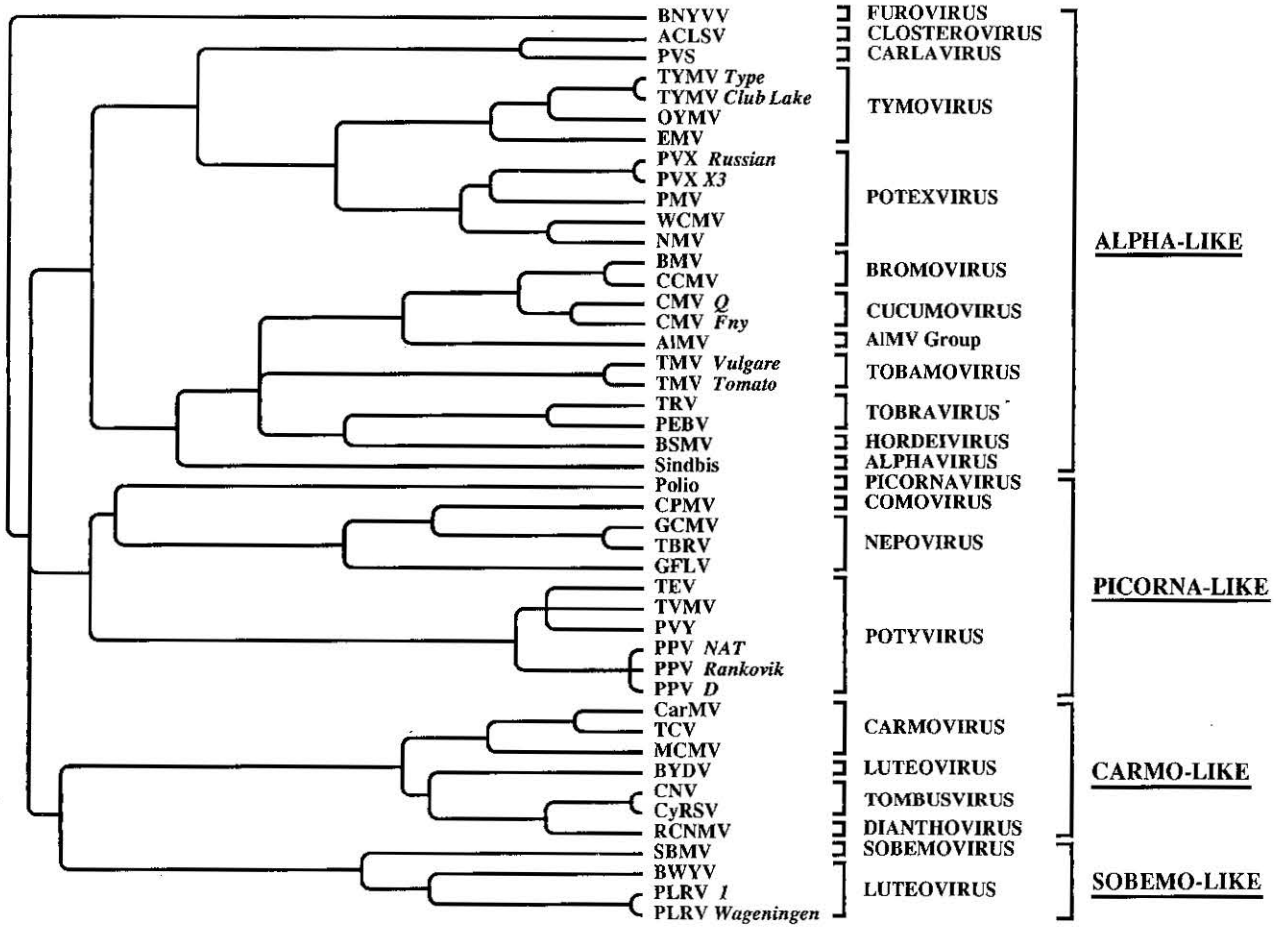

Figure 3. Clustering of plant RNA viruses using the conserved polymerase motifs. The degree of homology between sequences is inversely proportional to the length of the branches. Modified with permission after Candresse *et al.*[21]

Eighth International Congress of Virology in Berlin,[27] that only some strains of BYDV (e.g. PAV) specify a carmo-like polymerase while other strains of this virus (e.g. GPV) have a sobemo-like polymerase. Comparison of the polymerase sequences of all carmo-like plant viruses suggests that if recombination is the underlying event that has led to the insertion of a carmo-like polymerase gene in the BYDV-PAV genome, recombination has most likely occurred between a prototype BYDV and a dianthovirus.[27]

Conclusions

Plant positive-strand RNA viruses can be clustered into four supergroups, primarily based on amino acid sequence homologies in their polymerase domains (the only domain present in all positive-strand RNA viruses). This grouping is strengthened by other features of each of these groups, such as the presence or absence of other conserved domains in the proteins involved in replication, and structures at the ends of the genomic RNA (Table 1). Whereas the α-like and picorna-like supergroups are each closely related to an animal virus family (see also the article on picorna-like viruses in this issue), the carmo-like and sobemo-like supergroups do not seem to have such clear animal counterparts so far. Homology can be detected to some extent between the polymerase domains of the carmo-like and the animal flaviviruses, but they differ widely in many other aspects. Hence we feel reluctant to call these small plant RNA viruses 'flavi-like'. The sobemo-like viruses do not have any known animal virus counterpart and this supergroup shares some picorna-like as well as carmo-like characteristics.

Taking all interviral similarities and, more importantly, differences into account, the formerly defined[2] α-like supergroup has thus been split into three supergroups. Even the 'core' α-like supergroup might be further divided into two subgroups based on amino acid sequence homologies in the polymerase and NTB domains (refs 20,21; Figure 3) and on clear differences in the third (amino-terminal) conserved domain of the replication proteins (ref 19; Figure 1).

Acknowledgements

We are grateful to Dr T. Candresse for providing the dendrogram shown in Figure 3, to Dr E. Koonin for providing results prior to publication and to Ms A. M. van der Jagt for preparation of the manuscript.

References

1. Goldbach RW (1986) Molecular evolution of plant RNA viruses. Annu Rev Phytopathol 24:289-310
2. Goldbach RW, Wellink J (1988) Evolution of plus-strand RNA viruses. Intervirology 29:260-267
3. Strauss JH, Strauss EG (1988) Evolution of RNA viruses. Annu Rev Microbiol 42:657-683
4. Goldbach R (1990) Genome similarities between positive-strand RNA viruses from plants and animals, in New Aspects of Positive Strand RNA Viruses (Brinton MA, Heinz FX, eds), pp 3-11. ASM, Washington DC
5. Rochon DM, Tremaine JH (1989) Complete nucleotide sequence of the cucumber necrosis virus genome. Virology 169:251-259
6. Habili N, Symons RM (1989) Evolutionary relationship between luteoviruses and other RNA plant viruses based on sequence motifs in their putative RNA polymerases and nucleic acid helicases. Nucleic Acids Res 23:9543-9555
7. Koonin EV (1990) Towards a phylogenetic taxonomy of positive strand RNA viruses. 8th International Congress of Virology, Berlin, Abstr W52-005
8. Miller RM, Purcell RM (1990) Hepatitis C virus shares amino acid sequence similarity with pestiviruses and flaviviruses as well as members of two plant virus supergroups. Proc Natl Acad Sci USA 87:2057-2061
9. Hahn YS, Gralcoui A, Rice CM, Strauss EG, Strauss JH (1989) Mapping of RNA$^-$ temperature-sensitive mutants of Sindbis virus: complementation group F mutants have lesions in nsP4. J Virol 63:1194-1202
10. Hahn YS, Strauss EG, Strauss JH (1989) Mapping of RNA$^-$ temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B and G to non-structural proteins. J Virol 63:3142-3150
11. Kroner P, Richards D, Traynor P, Ahlquist P (1989) Defined mutations in a small region of the bromo mosaic virus 2a gene cause diverse temperature-sensitive RNA replication phenotypes. J Virol 63:5302-5309
12. Traynor P, Ahlquist P (1990) Use of bromovirus RNA2 hybrids to map *cis*- and *trans*-acting functions in a conserved RNA replication gene. J Virol 64:69-77
13. Kamer G, Argos P (1984) Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. Nucleic Acids Res 12:7269-7282
14. Gorbalenya AE, Blinov VM, Koonin EV (1985) Prediction of nucleotide-binding properties of virus-specific proteins from their primary structure. Mol Genet 11:30-36
15. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1988) A conserved NTP-motif in putative helicases. Nature 333:22
16. Gorbalenya AE, Koonin EV, Donchenko AP, Blinov VM (1989) 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
17. Hodgman TC (1988) A new superfamily of replicative proteins. Nature 333:22-23
18. Mi S, Durbin R, Huang HV, Rice CM, Stollar V (1989) Association of the Sindbis virus RNA methyltransferase activity with the nonstructural protein nsP1. Virology 170:385-391
19. Rozanov MN, Morozov SY, Skryabin KG (1990) Unexpected close relationship between the large nonvirion proteins of filamentous potexviruses and spherical tymoviruses. Virus Genes 3-4:373-379
20. Morozov SG, Dolja VV, Atabekov JG (1989) Probable reassortment of genomic elements among elongated RNA-containing plant viruses. J Mol Evol 29:52-62
21. Candresse T, Morch MD, Dunez J (1990) Multiple alignment and hierarchical clustering of conserved amino acid sequences in replication-associated proteins of plant RNA viruses. Res Virol 141:315-329
22. Hull R (1989) The movement of viruses in plants. Annu Rev Phytopathol 27:213-240
23. Gorbalenya AE, Koonin EV, Blinov VM, Donchenko AP (1988) Sobemovirus genome appears to encode a serine protease related to cysteine proteases of picornaviruses. FEBS Lett 236:287-290
24. Gorbalenya AE, Donchenko AP, Blinov VM, Koonin EV (1989) Cysteine proteases of positive strand RNA viruses and chymotrypsin-like serine proteases. FEBS Lett 243: 103-114
25. Veidt I, Lot H, Leiser M, Scheidecker D, Guilley H, Richard K, Jonard G (1988) Nucleotide sequence of beet western yellows virus RNA. Nucleic Acids Res 16:9917-9932
26. Bazan JF, Fletterick RJ (1989) Comparative analysis of viral cysteine protease structural models. FEBS Lett 249:5-7
27. Waterhouse P, Keese P, Young M, Martin B, Zhuomin C, Gerlach W (1990) Inter- and intra-group relationships of luteoviruses. 8th International Congress of Virology, Berlin, Abstr W52-004