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.

RNA viruses: genome structure and evolution

Ellen G. Strauss and James H. Strauss

California Institute of Technology, California, USA

The explosive pace of sequencing of RNA viruses is leading to rapid advances in our understanding of the evolution of these viruses and of the ways in which their genomes are organized and expressed. New insights are coming not only from genomic nucleotide sequence comparisons, but also from direct sequencing of transcribed mRNAs and of RNAs that serve as intermediates in replication.

Current Opinion in Genetics and Development 1991, **1**:485–493

Introduction

During the past decade, advances in the technology of cloning and sequencing have made possible the very rapid acquisition of information about the organization of RNA virus genomes. At the current time, complete genomic sequences exist for at least one representative of almost all the known RNA virus groups; this information has led to the elucidation of evolutionary relationships between many superficially diverse groups. It has often been found that overall strategies of replication, such as the relative location within a genome of genes of similar function, the presence or lack of subgenomic mRNAs, readthrough of termination codons for downstream products, or use of an 'ambisense' transcription strategy, show that certain viruses are distantly related, even when no sequence homology remains [1]. In addition, a number of short amino-acid-sequence elements have been recognized as indicators of the function of particular proteins and it has been suggested that proteins sharing such motifs are also related by descent from common ancestors. These indicators include the 'GDD motif', which is characteristic of RNA polymerases [2,3,4•], the 'DEAD motif' [5] and the 'G-x-x-GKS/T motif', which are characteristic of RNA helicases [5–7], and the sequence elements surrounding the amino acids in the active sites of viral-encoded proteases [8].

Even among well studied viruses known to be closely related, sequencing often clarifies their relationships. From sequence analysis, the genera of the *Picornaviridae* have been realigned [9] and the taxonomy of the *Paramyxoviridae* clarified (see discussion below). Sequence comparisons have also been useful in epidemiological studies, both longitudinal studies to determine rates of change of a particular virus in nature, and comparisons of geographic isolates to pinpoint the origins of epidemic viruses [10–12,13•].

One of the most exciting and unanticipated results of genome structure comparisons has been the discovery that plant virus counterparts exist for almost every major group of RNA-containing animal viruses. In the case of the bunyaviruses, the plant and animal representatives are so closely related as to be placed within the same family, and an argument could be made that the plant tenuiviruses should be considered as belonging to the same genus as the animal uukuviruses and phleboviruses. In other cases, the plant and animal viruses are counterparts in the sense that they share genome organization and transcription strategies, and may share sequence homology in a number of proteins (even though each may possess unique genes required for replication in their respective hosts), but the plant and animal virions may be quite different in morphology. The existence of such viruses indicates that plant and animal viruses have radiated from a small number of ancestral prototypic viruses, and that the repertoire of successful replication modes may be limited. These studies have also made it clear that recombination has played an important role in the evolution of RNA viruses, and that viruses can acquire the ability to jump across wide phylogenetic barriers, whether by recombination or adaptation, rather more easily than would have been suspected a decade ago.

Because it is impossible to discuss all of the significant advances of the past year in a short review, we have selected three areas of particular interest. These include: a newly described mechanism used by the paramyxoviruses to shift the translation frame that is useful for the classification of these viruses; updated information

Abbreviations

BSMV—barley stripe mosaic virus; **L**—large; **Ldr**—leader RNA; **M**—medium; **NDV**—Newcastle disease virus; **NTR**—non-translated region; **ORF**—open reading frame; **PIV**—parainfluenza virus; **S**—small; **ssRNA**—single-strand RNA; **SV5**—simian virus 5; **TMV**—tobacco mosaic virus; **vc RNA**—virus complementary RNA.

 ISSN 0959-437X

on the replication strategy of coronaviruses; and additional insights into plant virus counterparts of animal viruses.

Coding strategy of the V/P genes of the *Paramyxoviridae*: a new mechanism for translational frame-shifting

Many viruses are known to increase the information content of their genomes by translating their RNA in more than one reading frame. Differential splicing (for viruses that replicate in the nucleus), translation initiation at more than one start codon, and ribosomal frame shifting have been described. In the V/P gene of the paramyxovirus simian virus 5 (SV5), two non-templated nucleotides are added during transcription of mRNAs to shift the reading frame in some, but not all, transcripts [14]. Within the past year, reports of similar mechanisms used by several other members of the paramyxovirus family have been described. Not only is the addition of non-templated G residues to shift the reading frame found in many, but not all, of these viruses, but the details differ among the different viruses and appear to be useful in the classification of the members of this family.

All members of the family *Paramyxoviridae*, in the order Mononegavirales, harbor a genome comprising single segment of negative polarity RNA (approximately 10–12 kb). The family contains three currently recognized genera: *pneumovirus*, *paramyxovirus*, and *morbillivirus*. The genus *paramyxovirus* includes Newcastle disease virus (NDV), the type virus, Sendai virus, human parainfluenza virus (PIV)-1, PIV-2, PIV-3, and PIV-4, mumps virus, and SV5, all of which contain a neuraminidase activity. The morbillivirus genus encompasses measles virus, rinderpest virus, and canine distemper virus, all of which are quite similar to paramyxoviruses but lack the neuraminidase activity. Pneumoviruses (respiratory syncytial virus and pneumonia virus of mice) are distinct, and contain a number of extra genes in addition to the nucleocapsid (N), phosphoprotein (P), membrane protein (M), glycoprotein (G), fusion protein (F), and large polymerase protein (L) genes common to all members of the family. All of the genes are monocistronic with the exception of the P (or V/P) genes.

Paramyxoviruses and morbilliviruses, but not pneumoviruses, increase the coding capacity of the V/P gene by translating products from more than one reading frame. The details of the V/P gene strategy are illustrated in figure 1. Two mechanisms are used: initiation of translation at two different AUGs, and the addition of non-templated G residues to shift the reading frame. In some cases all three reading frames are utilized and up to four protein products are produced from the V/P gene. In measles [15], PIV-1 [16•], PIV-3 [17••], and Sendai viruses [18], a P protein of approximately 600 amino acids is translated from an mRNA that is a faithful complement of the genome. In addition, a smaller protein (C) of about 200 amino acids is translated from the same mRNA by internal initiation at a methionine in a second reading frame. The V protein, which is amino-coterminal with P but contains a carboxy-terminal domain that is highly conserved and rich in cysteine residues, is encoded by an mRNA that is formed when a single non-templated G residue is inserted into the mRNA during transcription. This shifts the frame to the third possible frame, resulting in a V protein of about 400 amino acids. Note that for PIV-3, an mRNA containing two non-templated G residues is also produced and is translated into a fourth protein, the D protein.

In PIV-2 [19•], PIV-4 [20•], SV5 [14], mumps [21•], and probably NDV, there is no single long open reading frame (ORF). Instead there are two significant ORFs which overlap in the middle of the gene. Translation of the mRNA resulting from faithful copying of the genome gives rise to the V protein. In the case of PIV-2 and SV5, exactly two non-templated G residues are inserted to shift the frame to the P protein frame; in PIV-4 the number of G's inserted is more variable, although the specific insertion of two gives the P protein. For mumps, insertion of two G's results in the mRNA for the P protein, whereas insertion of four G's gives the message for another product, the I protein (Fig. 1). All of the P proteins of this latter group are 391–399 amino acids long.

The functions of all of these proteins in virus replication have not yet been established, nevertheless it is fascinating that in viruses that are so closely related and similar in many aspects of their replication, some should use one gene to translate only one protein, whereas up to four proteins are produced from a single gene in others. The disparity in use of the P protein among the *Paramyxoviridae* also suggests either that evolution to use multiple reading frames within a gene proceeds more rapidly than might have been predicted some time ago, or that the ancestral paramyxovirus used multiple reading frames, and that as the function of some of these translation products became non-essential during evolutionary divergence some paramyxoviruses lost the ability to produce them.

Based on this information, there has been a suggestion to reclassify the *Paramyxoviridae* [20•]. The morbilliviruses and pneumoviruses would remain unchanged, but the paramyxovirus genus would be split into two genera. One genus would contain PIV-1, PIV-3, and Sendai virus, and the second would contain PIV-2, PIV-4, SV5, mumps, and NDV. Notably, a recent analysis of the aligned sequences of the L protein of these various viruses led to precisely the same assignment [22••].

New aspects of coronavirus replication

Many RNA viruses are known to produce subgenomic mRNAs. The plus-stranded togaviruses as well as a number of plus-stranded plant viruses transcribe a subgenomic mRNA for the structural protein(s), and all of

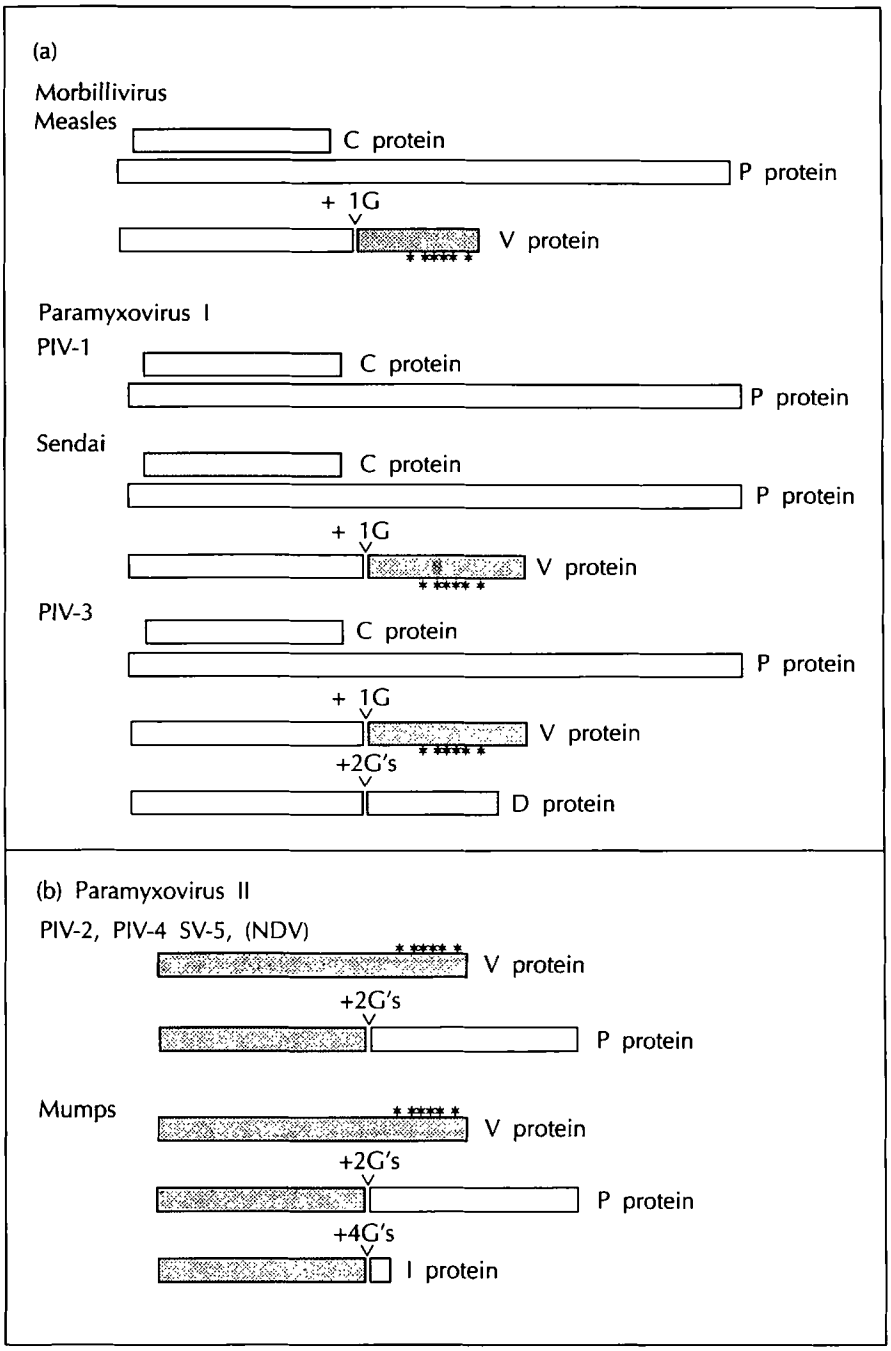

Fig. 1. Transcription and translation strategy of the V/P genes of the paramyxoviruses. In all cases the three types of shading represent the three possible open reading frames (ORFs) and the cysteine-rich domain is indicated by the asterisks. (a) For the morbillivirus group and the paramyxovirus group I the top two boxes are protein products translated in different frames from an mRNA exactly complementary to the genomic RNA. The third ORF, accessed by the addition of a single non-templated nucleotide during RNA transcription, is shown with dark gray shading. For PIV-3 the addition of two non-templated nucleotides produces a protein with a carboxy-terminal domain in the same frame as the C protein. (b) For paramyxovirus group II the product of the exactly complementary mRNA is the V protein containing the cysteine-rich domain (dark gray shading and asterisks) and the addition of the non-templated nucleotides shifts the frame to produce the P protein. References from which this information was obtained are cited in the text.

the minus-stranded viruses produce mRNAs that are less than genome length. These mRNAs lack a *cis*-acting terminal element (either a sequence or a structure) essential for initiation of replication by the viral replicase, and thus cannot replicate in infected cells. The coronaviruses also produce up to seven different subgenomic mRNAs of genomic polarity. Sethna *et al.* [23] reported the presence of minus-strand copies of these mRNAs in coronavirus-infected cells and suggested that the subgenomic RNAs could replicate. As discussed below, during the past year a number of groups have reported that coronavirus mRNAs are indeed actively replicated in infected cells. The *Coronaviridae* are plus-stranded RNA viruses with the largest known RNA genomes, more than 27 kb. The genome organization of a typical coronavirus is illustrated in figure 2a (reviewed in [24]). Almost 20 kb at the 5′ end is devoted to two long ORFs encoding the RNA replicase. The replicase is translated from the genomic RNA as two large polyproteins, with the larger, presumably produced by ribosomal frame-shifting, corresponding to the entire 20 kb region. These polyproteins are believed to be posttranslationally cleaved by two cysteine proteases, which are encoded within them, into an unknown number of final products [25,26]. A nested set of six or seven 3′ coterminal subgenomic mRNAs encode the virion proteins and a variable number of non-structural components. The structural proteins include a nucleo capsid protein, N, that packages the RNA into a heli-

cal nucleocapsid, and two virus-encoded membrane glycoproteins, M and S, present in a lipid containing envelope. The M protein is O-glycosylated, very hydrophobic, and is thought to traverse the membrane several times. The S protein is N-glycosylated and forms projections external to the lipid bilayer (reviewed in [24]). In addition, certain coronaviruses contain an additional envelope glycoprotein that has significant homology with the hemagglutinin of influenza C virus, and that possesses demonstrable hemagglutinating and esterase activities [27].

Each mRNA contains an identical 5′ leader sequence derived from the 5′ end of the genome and it is widely accepted that these RNAs are transcribed by a mechanism of 'leader-primed' transcription from the full length minus-strand complement of the genome [28]. In this model, the polymerase first copies the leader sequence, dissociates from the template, and subsequently reassociates with the same or a different template to reinitiate transcription from internal sites. Such a strategy could explain the high frequency of recombination observed for coronaviruses. At first, it was thought that this was the only mechanism for production of the subgenomic mRNAs. Recently, however, evidence has accumulated that minus-strand RNAs complementary to each of the mRNAs exist in infected cells, that these are active in transcription, and that such replication of the mRNAs leads to the formation of a significant fraction of the mRNA present in the infected cell [23,29,30•,31••]. These minus strands are exact complements of the mRNAs and contain anti-leader sequences (Fig. 2b) [31••].

Two related enteric Toroviruses (Breda virus and Berne virus), with an unusual nucleocapsid morphology, have been identified that share characteristics with the coronaviruses [32••]. Toroviruses have the same basic gene order as the coronaviruses, 5′–polymerase–spike-protein–M-protein–capsid–3′, and translate these polypeptides from a nested set of four 3′coterminal mRNAs. There are regions of amino acid similarity within the polymerase genes of the two groups. Moreover, toroviruses contain an ORF which encodes part of the influenza C-like hemagglutinin mentioned above. The function of this gene is obscure and it is unclear whether toroviruses and coronaviruses independently obtained this gene via recombination, or whether during evolution the gene was

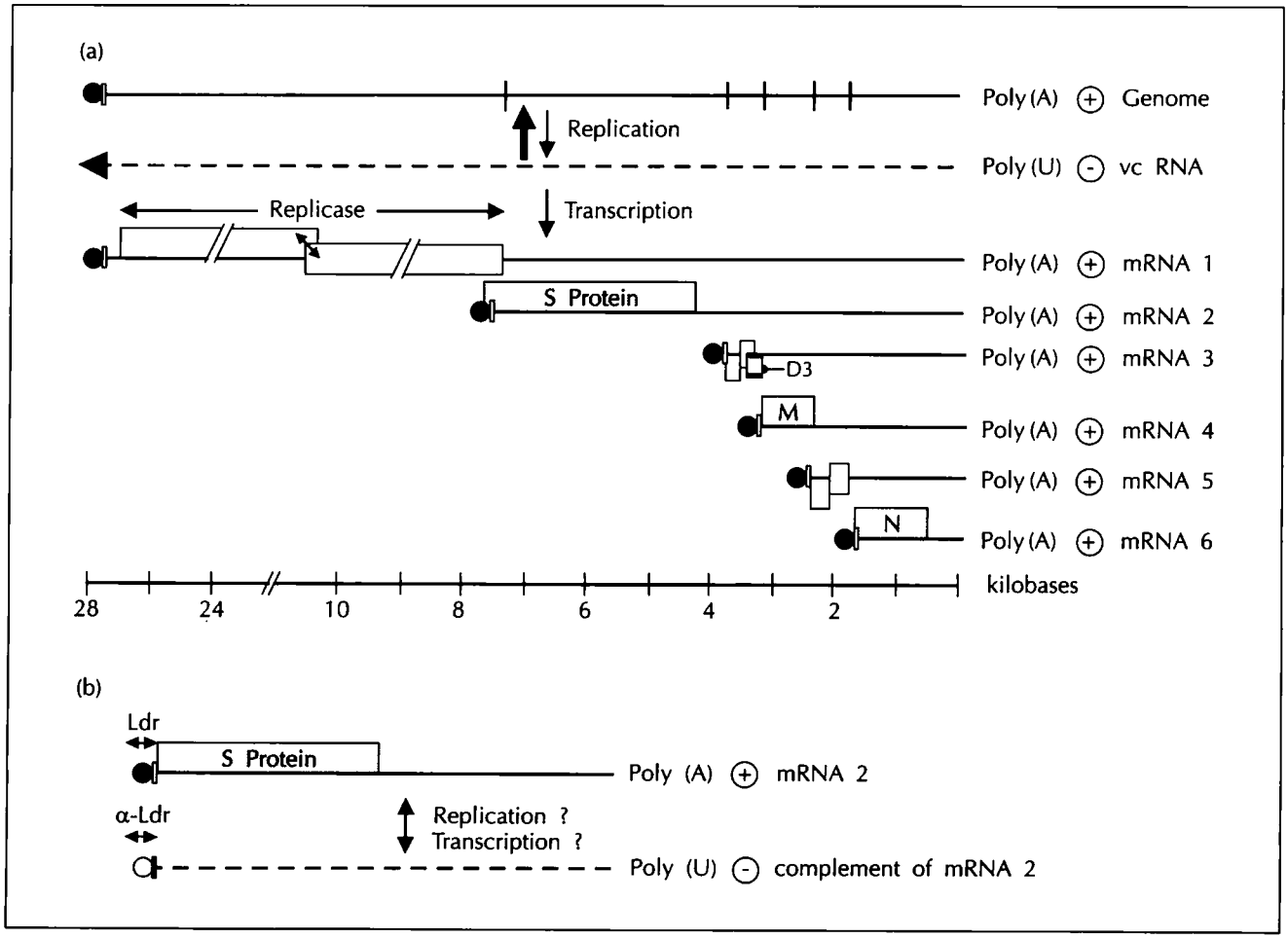

Fig. 2. Replication strategy of a typical coronavirus. (a) The dark ball at the 5′ end of the genome is the leader sequence. The labelled boxes are the unique open reading frames (ORFs) of each subgenomic mRNA that are read; note that in mRNA 3 it is the third ORF (5′ to 3′) in the unique domain that is preferentially expressed as D3. mRNA 1 is identical to the genomic RNA. The mRNAs are drawn to scale, and virion structural proteins are shaded. (b) Structure of a minus-strand copy of a typical mRNA. Ldr, leader RNA; α-Ldr, complement of the leader RNA.

lost from some coronavirus-like viruses but not from others. All of the shared characteristics suggest that the toroviruses and coronaviruses have diverged from a common ancestor, but this argument will be considerably strengthened if it is found that toroviruses, like coronaviruses, have independently replicating mRNAs.

Plant viruses related to animal viruses

One of the more interesting discoveries of the past decade has been the finding that many plant viruses have animal virus counterparts (Table 1) to which they are related to varying degrees. This is a topic that has received a great deal of attention in the last year, as more complete sequences of both plant and animal viruses have been determined. A recent volume of *Seminars in Virology* was devoted exclusively to this topic [33•].

The relationships between the animal picornaviruses and the plant comoviruses and between the animal togaviruses and the plant tobamoviruses, bromoviruses, and alfalfa mosaic virus were first described several years ago [34–37]. These relationships involve similarities in genome organization and clear amino-acid-sequence homology in some, but not all, of the encoded proteins, suggesting quite strongly that the 'picorna-like plant viruses' and *Picornaviridae* are evolutionarily related to one another as are 'Sindbis-like plant viruses' and *Togaviridae*.

The genomes of Sindbis virus, the type alphavirus, family *Togaviridae*, and tobacco mosaic virus (TMV), are compared in figure 3. The relationship between Sindbis virus and TMV is straightforward. In each case the replicase genes are translated from the genomic RNA, readthrough of a termination codon is required to translate the RNA polymerase, and there are long stretches of clear amino acid sequence homology in three genes. The structural proteins are translated from a subgenomic mRNA and appear to be unrelated to one another. TMV is a rod-shaped virus, and Sindbis virus is enveloped. The recent demonstration that the nucleocapsid protein of Sindbis virus is structurally related to chymotrypsin is an exciting development and suggests that this capsid may have been obtained by recombination from a cellular protease (Rossman, personal communication). The RNA replication signals in the 3′ non-translated region (NTR) are also different in the two viruses. The Sindbis virus 3′ NTR contains a number of sequence elements, including repeated sequences and a 19 nucleotide conserved element, that are believed to function as linear elements, and which terminate in a poly(A) tract. TMV contains a 3′terminal nucleotide sequence capable of forming a tertiary structure similar to that of tRNA that is recognized by an aminoacyl tRNA synthetase. These tRNA-like sequences in TMV and a number of other plant virus RNAs are thought to be important for RNA replication and are known to be essential for infectivity, because certain point mutations within the sequence are lethal [38]. Upstream of this structure are a number of stem and loop elements in which certain bases in the loops can hydrogen bond with sequences adjacent to the stems, forming 'pseudo-

Table 1. Similar single-strand RNA viruses infecting plants and animals.

	Animal viruses			Plant viruses		
Genome type	Virus family	Morphology	Example	Virus group	Morphology	Example
ssRNA–non-segmented plus-strand, small	*Picornaviridae*	Icosahedral, non-enveloped	poliovirus	Comovirus	Icosahedral, non-enveloped	CMV
Plus-strand, medium	*Togaviridae*	Icosahedral, enveloped	Sindbis	Tobamovirus Bromovirus	Helical, rod Icosahedral	TMV BMV
Minus-strand, small	*Rhabdoviridae*	Bullet-shaped, enveloped	VSV	Phytorhabdovirus	Bullet-shaped, eveloped	SYNV
ssRNA–segmented minus-strand	*Bunyaviridae* Phlebovirus	Enveloped, helical	Rift Valley	Tospovirus	Enveloped, helical nucleocapsid	TSWV
	Uukuvirus	Nucleocapsid	Uukuniemi	Tenuivirus	Flexible rod	RStV

BMV, brome mosaic virus; CMV, cowpea mosaic virus; RSt V, rice stripe virus; SYNV, sonchus yellow net virus; TMV, tobacco mosaic virus; TSWV, tomato spotted wilt virus; VSV, vesicular stomatitis virus.

knots' [39]. Recently, it has been shown that the 3′ NTR of TMV (containing both the tRNA and pseudoknot domains) can substitute functionally for a poly(A) tail in the expression of heterologous mRNAs in both plant and animal cells [40••]. Some synergistic interaction with other viral elements is suggested by the fact that maximal expression occurs in certain constructs in which the 5′ NTR and 3′ NTR of TMV flank the heterologous reporter gene [40••].

Members of at least five other groups of plant viruses belong to the 'Sindbis-like superfamily', although in some cases the RNA genomes are divided into multiple independent RNA segments, as in the case of bromoviruses and hordeiviruses (Fig. 3). Barley stripe mosaic virus (BSMV), a hordeivirus, has now been sequenced and found to have a number of interesting properties that illustrate the evolution of these viruses [41•]. In BSMV, the three domains of sequence similarity to Sindbis virus are encoded on separate gene segments. Remarkably, homologs to the helicase domain found in the Sindbis virus non-structural protein, nsP2, are present on BSMV RNA segments, although only one is required for RNA replication [41•]. Two notable characteristics of hordeiviruses are: that the 3′ NTRs combine several motifs from 3′ to 5′, that is, a tRNA structure, a number of psuedoknots and finally a poly (A) tract adjacent to the end of the ORF (reviewed in [42]); and that a certain amount of plasticity is apparent in the polymerase, encoded on RNA 3. Three different forms of RNA 3 were originally identified in different strains of BSMV, although it has now been shown that all three can occur together in certain strains. Form IV contains a deletion in an essential polymerase domain, and is therefore defective and rapidly removed from the population. However, both Form III and Form II, which contain a tandem duplication of 350–370 nucleotides at the amino-terminus of the ORF for the polymerase, often occur [41•]. Comparable variation is not seen within the second ORF of RNA 3 which encodes a cysteine-rich polypeptide necessary for systemic infection of plants. Thus, a remarkable series of recombination events were involved in the production of the BSMV genome, and the virus may still be evolving toward an optimal organization.

The discovery of plant viruses with segmented negative-strand genomes that are closely related to the animal viruses in the family *Bunyaviridae* is also of considerable interest for viral taxonomy. The *Bunyaviridae* have three genome segments of negative-strand RNA. The largest segment encodes the RNA replicase, the medium sized segment encodes the glycoproteins found in the viral envelope, and the smallest segment encodes the nucleocapsid protein. Five genera of *Bunyaviridae* have been described, which differ from one another in details of their genome organization and sequence, and in their vectors. Nairoviruses and hantaviruses translate a single polypeptide from a single ORF in the small segment; bunyaviruses

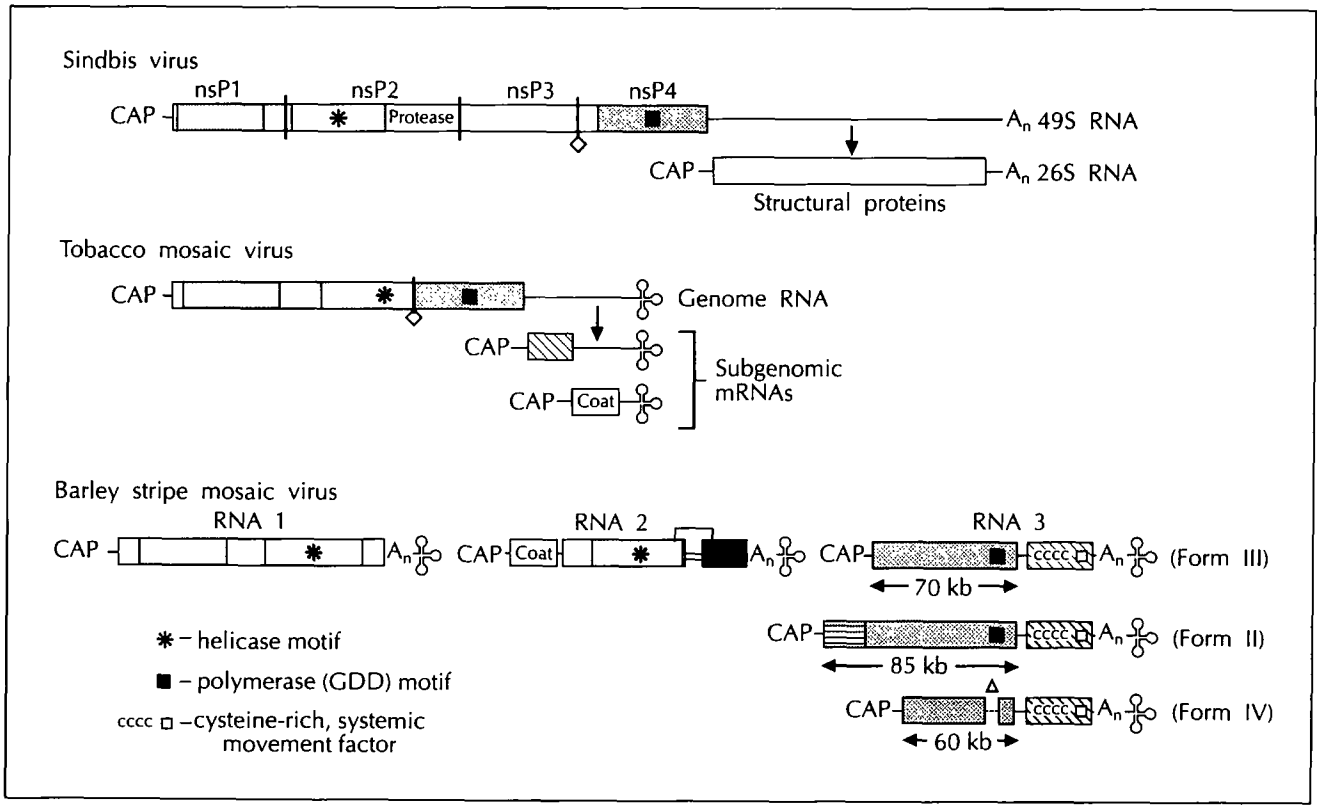

Fig. 3. Comparison of the replication strategies of Sindbis virus with tobacco mosaic virus and barley stripe mosaic virus. All domains shaded in the same way have significant amino acid sequence similarity to one another. The open diamond is the termination codon readthrough to produce the polymerase. The Sindbis non-structural proteins nsP1, nsP2, nsP3, and nsP4 are responsible for RNA replication. Three short sequence motifs are indicated with symbols (see key).

translate two polypeptides from overlapping ORFs in the small segment by initiation at two different AUG codons; and phlebovirus and uukuvirus small segments have two non-overlapping ORFs that are read in opposite orientations (i.e. one is translated from a genome sense mRNA and one from an antigenome sense mRNA) [43]. This strategy, termed 'ambisense', is also found in the bipartite *Arenaviridae*. Virions of *Bunyaviridae* with ambisense small segments contain small segment RNA of both polarities (although not in equal concentrations), but their medium segment RNA is all negative strand.

The tospoviruses (for **to**mato **spo**tted wilt virus), with a morphology very similar to bunyaviruses, have three genome segments of minus-strand RNA, an ambisense translation strategy for the s segment [44••], and have been classified as a genus of *Bunyaviridae*. The taxonomic position of rice stripe virus, the type tenuivirus, is less clear. The virion is poorly defined and the infectious material seems to consist of long thin filaments (perhaps circular) containing the polymerase, which are reminiscent of bunyavirus nucleocapsids. The RNA consists of four segments, of which both segment 3, encoding the capsid protein, and segment 4, of unknown function, are ambisense [45•,46••]. The results of Kakutani *et al.* [46••] describe the first example of a virus with more than one ambisense segment. Unfortunately, no sequence information is available on the medium or large segment RNAs, which would be expected to show the highest similarity to the comparable RNA segments of *Bunyaviridae*.

At the 3′ termini of the three *Bunyaviridae* RNA segments there are short nucleotide sequences (10–12 residues in length) that are conserved within a given genus and that are complementary to conserved sequences at the 5′ termini such that the RNAs can form panhandles. It has been suggested that these sequences are promoter elements for the initiation of RNA synthesis, and that in a mixed infection only *Bunyaviridae* with identical or nearly identical termini would be replicated by the same polymerase. This would imply that only members of the same genus can exchange genome segments. Nine out of 10 of the terminal nucleotides are identical for phleboviruses and uukuviruses. In addition, sequence similarities have been found in the amino acid sequences of the N proteins and Gl and G2 glycoproteins of these two genera. Surprisingly, the terminal sequence for the tenuiviruses is 90% identical to uukuviruses, suggesting that these two groups are very closely related, but the termini of tospoviruses bear no relationship to the termini of any other bunyavirus. All of these characteristics lead us to suggest that, despite their ambisense strategy, the tospoviruses have diverged significantly from other *Bunyaviridae*, in contrast to tenuiviruses, which may be very closely related to *Bunyaviridae*, despite their distinct morphology. Furthermore, we propose that the viral phleboviruses and uukuviruses belong to a single genus.

Conclusions

During the past year, a number of interesting insights into the interrelationships among RNA viruses have been obtained. There appear to be many similarities among the replication strategies of seemingly diverse viruses; representatives of plant virus groups and animal virus families often share certain genes and features of genome organization while differing in other aspects of their replication. As more viral genomic sequences are obtained, a clearer picture of the RNA virus phylogenetic tree emerges, and it appears that RNA viruses extant today have evolved from a small number of protoviruses. In addition to divergent evolution, RNA viruses also evolve by recombination and the acquisition of new genes either from other viruses or from their hosts. The net result of such a reshuffling of entire viral genes is a form of modular evolution, where segments of the genome are transferred as a module.

In recent years, a number of new strategies of viral gene expression have been discovered. Because most RNA virus genomes are small, perhaps limited by the inherent error frequency caused by RNA replication without proofreading, RNA viruses are quite efficient, and have evolved a range of mechanisms by which to expand the available coding capacity and differentially regulate their gene products. These include translating more than one reading frame starting at different initiation sites, differential splicing of mRNAs during transcription, and production of multiple mRNAs in which the reading frame has been shifted by the insertion of non-templated nucleotides. Furthermore, in at least one system subgenomic mRNAs can replicate independently, and this replication may be an additional important mechanism for regulating the amounts of individual gene products.

Acknowledgements

Work of the authors is supported by grants AI10793 and AI20612 from the National Institutes of Health and by grant DMB 9104054 from the National Science Foundation.

References and recommended reading

Papers of special interest, published within the annual period of review, have been highlighted as:

- • of interest
- •• of outstanding interest

1. STRAUSS EG, STRAUSS JH: **Replication Strategies of the Single Stranded RNA Viruses of Eukaryotes.** *Curr Topics Microbiol Immunol* 1983, 105:1–98.

2. KAMER G, ARGOS P: **Primary Structural Comparison of RNA-dependent Polymerases from Plant, Animal and Bacterial Viruses.** *Nucleic Acids Res* 1984, 12:7269–7282.

3. ARGOS P: **A Sequence Motif in Many Polymerases.** *Nucleic Acids Res* 1988, 16:9909–9916.

4. • BRUENN JA: **Relationships Among the Positive Strand and Double-strand RNA Viruses as Viewed Through their**

RNA-dependent RNA Polymerases. *Nucleic Acids Res* 1991, 19:217–226.
Sequences of 50 RNA-dependent RNA polymerases from 43 plus-strand viruses and seven double-strand viruses have been aligned, and phylogenetic trees constructed. A large number of these viruses form a single cluster whose only common characteristic is insect transmission.

5. GORBALENYA AE, KOONIN EV, DONCHENKO AP, BLINOV VM: **Two Related Superfamilies of Putative Helicases Involved in Replication, Repair, and Expression of DNA and RNA Genomes.** *Nucleic Acids Res* 1989, 17:4713–4730.

6. ZIMMERN D: **Evolution of RNA Viruses.** In *RNA Genetics* edited by Domingo E, Holland JJ, Ahlquist P [book]. Boca Raton Florida: CRC Press, 1988, pp 211–240.

7. GORBALENYA AE, KOONIN EV, DONCHENKO AP, BLINOV VM: **A Conserved nTP-Binding Motif in Putative Helicases.** *Nature* 1988, 333:22.

8. HARDY WR, STRAUSS JH: **Processing the Nonstructural Proteins of Sindbis Virus: Nonstructural Proteinase is in the C-Terminal Half of nsP2 and Functions Both in *Cis* and *Trans*.** *J Virol* 1989, 63:4653–4664.

9. PALMENBERG AC: **Sequence Alignments of Picornaviral Capsid Proteins.** In *Molecular Aspects of Picornavirus Infection and Detection* edited by Semler B, Ehrenfeld E [book]. Washington, DC: American Society for Microbiology, 1989, pp 211–241.

10. KING AMQ, UNDERWOOD BO, MCCAHON D, NEWMAN JWI, BROWN F: **Biochemical Identification of Viruses Causing the 1981 Outbreaks of Foot-and-Mouth Disease in the UK.** *Nature* 1981, 293:479–480.

11. TAKEDA N, MIYAMURA K, OGINO T, NATORI K, YAMAZAKI S, SAKURAI N, NAKAZONO N, ISHII K, KONO R: **Evolution of Enterovirus Type 70: Oligonucleotide Mapping Analysis of RNA Genome.** *Virology* 1984, 134:375–388.

12. BURNESS AT, PARDOE I, FARAGHER SG, VRATI S, DALGARNO L: **Genetic Stability of Ross River Virus During Epidemic Spread in Nonimmune Humans.** *Virology* 1988, 167:639-643.

13. • SHIRAKO Y, NIKLASSON B, DALRYMPLE JM, STRAUSS EG, STRAUSS JH: **Structure of the Ockelbo Virus Genome and its Relationship to other Sindbis Viruses.** *Virology* 1991, 182:753–764.
Complete sequence of the genome of Ockelbo virus, a strain of Sindbis virus which causes epidemic disease in northern Europe. From sequence comparisons it was found that Ockelbo virus must have been transferred to northern Europe from South Africa.

14. THOMAS SM, LAMB RA, PATERSON RG: **Two mRNAs That Differ by Two Nontemplated Nucleotides Encode the Amino Coterminal Proteins P and V of the Paramyxovirus SV5.** *Cell* 1988, 54:891–902.

15. CATTANEO R, KAELIN K, BACZKO K, BILLETER MA: **Measles Virus Editing Provides an Additional Cysteine-rich Protein.** *Cell* 1989, 56:759–764.

16. • MATSUOKA Y, CURRAN J, PELET T, KOLAKOVSKY D, RAY R, COMPANS RW: **The P gene of Human Parainfluenza Virus Type 1 Encodes P and C Proteins But not a Cysteine-rich V Protein.** *J Virol* 1991, 65:3406–3410.
Demonstration that PIV-1 does not need a cysteine-rich V protein, and that the addition of non-templated G residues is not essential for all paramyxo-viruses.

17. •• PELET T, CURRAN J, KOLAKOVSKY D: **The P Gene of Bovine Parainfluenza Virus 3 Expresses all Three Reading Frames From a Single mRNA Editing Site.** *EMBO J* 1991, 10:443–448.
A comprehensive look at all of the mRNA transcripts and translated protein products of the P gene of PIV-3, illustrating that all three reading frames are used over a stretch of more than 300 nucleotides. This may be the first gene sequenced in which all three frames are used.

18. VIDAL S, CURRAN J, KOLAKOVSKY D: **Editing of the Sendai Virus P/C mRNA by G Insertion Occurs During mRNA Synthesis via a Virus-encoded Activity.** *J Virol* 1990, 64:239–246.

19. • SOUTHERN JA, PRECIOUS B, RANDALL RE: **Two Nontemplated Nucleotide Additions are Required to Generate the P mRNA of Parainfluenza Virus Type 2 Since the RNA Genome Encodes Protein V.** *Virology* 1990, 177:388–390.
To study the structure of the mRNAs from the PV gene, both genomic and mRNA sequences were amplified by PCR and sequenced.

20. • KONDO K, BANDO H, TSURUDOME M, KAWANO M, NISHIO M, ITO Y: **Sequence Analysis of the Phosphoprotein (P) Genes of Human Parainfluenza Type 4A and 4B Viruses and RNA Editing at Transcript of the P Genes: the Number of G Residues Added is Imprecise.** *Virology* 1990, 178:321–326.
Proof that the number of G residues added to mRNAs of the P gene of PIV-4 is imprecise. The P gene products for nine paramyxoviruses are aligned to provide a phylogenetic dendrogram.

21. • PATERSON RG, LAMB RA: **RNA Editing by G-Nucleotide Insertion in Mumps Virus P-Gene mRNA Transcripts.** *J Virol* 1990, 64:4137–4145.
Demonstration that the unedited transcript encodes the cysteine-rich V protein and that variable numbers of inserted G's give transcripts for two other products.

22. •• STEC DA, HILL MAI, COLLINS PL: **Sequence Analysis of the Polymerase L Gene of Human Respiratory Syncytial Virus and Predicted Phylogeny of Nonsegmented Negative-strand Viruses.** *Virology* 1991, 183:273–287.
The complete sequence of the polymerase gene of RSV is presented and compared with the L genes of *Rhabdoviridae* and *Paramyxoviridae*. Comprehensive phylogenetic trees are shown.

23. SETHNA PB, HUNG S-L, BRIAN DA: **Coronavirus Subgenomic Minus-strand RNAs and the Potential for mRNA Replicons.** *Proc Natl Acad Sci USA* 1989, 86:5626–5630.

24. LAI MMC: **Coronavirus: Organization, Replication and Expression of Genome.** *Annu Rev Microbiol* 1990, 44:303–333.

25. GORBALENYA AE, KOONIN EV, DONCHENKO AP, BLINOV VM: **Coronavirus Genome: Prediction of Putative Functional Domains in the Non-structural Polyprotein by Comparative Amino Acid Sequence Analysis.** *Nucleic Acids Res* 1989, 17:4847–4861.

26. BAKER SC, SHIEH C-K, SOE LH, CHANG M-F, VANNIER DM, LAI MMC: **Identification of a Domain Required for Autoproteolytic Cleavage of Murine Coronavirus Gene A Polyprotein.** *J Virol* 1989, 63:3693–3699.

27. VLASAK R, LUYTJES W, LEIDER W, SPAAN W, PALESE P: **The E3 Protein of Bovine Coronavirus is a Receptor-destroying Enzyme With Acetyl-esterase Activity.** *J Virol* 1988, 62:4686–4690.

28. LAI MMC: **Coronavirus Leader-RNA-primed Transcription: an Alternative Mechanism to RNA Splicing.** *Bioessays* 1986, 5:257–260.

29. SAWICKI SG, SAWICKI DL: **Coronavirus Transcription: Subgenomic Mouse Hepatitis Virus Replicative Intermediates Function in RNA Synthesis.** *J Virol* 1990, 64:1050–1056.

30. • HOFMANN MA, SETHNA PB, BRIAN DA: **Bovine Coronavirus mRNA Replication Continues Throughout Persistent Infection in Cell Culture.** *J Virol* 1990, 64:4108–4114.
Quantitation of the amount of minus strands corresponding to the nested set of mRNAs made during persistent infection. Suggests that replication of mRNAs is not a rare or unimportant event.

31. •• SETHNA PB, HOFMANN MA, BRIAN DA: **Minus-strand Copies of Replicating Coronavirus mRNAs Contain Antileaders.** *J Virol* 1991, 65:320–325.
Direct sequencing of minus strands corresponding to coronavirus mRNAs to show that they are exactly complementary to the mRNAs.

32. •• SNIJDER EJ, DEN BOON JA, HORZINEK MC, SPAAN WJM: **Comparison of the Genome Organization of Toro- and**

Coronaviruses: Evidence for Two Nonhomologous RNA Recombination Events During Berne Virus Evolution. *Virology* 1991, 180:448–452.
Sequence analysis to suggest that coronaviruses and toroviruses are related and that recombination has played a role in their evolution.

33. • MAHY BWJ (ED): **Related Viruses of Plants and Animals.** In *Seminars in Virology* 1991, 2:1–77.
A collection of brief reviews that summarizes the know facts about related plant and animal viruses.

34. HASELOFF J, GOELET P, ZIMMERN D, AHLQUIST P, DASGLUTA R, KAESBERG P: **Striking Similarities in Amino Acid Sequence Among Nonstructural Proteins Encoded by RNA Viruses that have Dissimilar Genomic Organization.** *Proc Natl Acad Sci USA* 1984, 81:4358–4362.

35. AHLQUIST P, STRAUSS EG, RICE CM, STRAUSS JH, HASELOFF J, ZIMMERN D: **Sindbis Virus Proteins nsP1 and nsP2 Contain Homology to Nonstructural Proteins from Several RNA Plant Viruses.** *J Virol* 1985, 53:536–542.

36. FRANSSEN H, LEUNISSEN J, GOLDBACH R, LOMONOSSOFF G, ZIMMERN D: **Homologous Sequences in Non-structural Proteins from Cowpea Mosaic Virus and Picornavirus.** *EMBO J* 1984, 3:855–861.

37. ARGOS P, KAMER G, NICKLIN MJH, WIMMER E: **Similarity in Gene Organization and Homology Between Protein of Animal Picornaviruses and a Plant Comovirus Suggest Common Ancestry of these Virus Families.** *Nucleic Acids Res* 1984, 12:7251–7267.

38. DREHER TW, HALL TC: **Mutational Analysis of the Sequence and Structural Requirements in Brome Mosaic Virus RNA for Minus Strand Promoter Activity.** *J Mol Biol* 1988, 201:31–40.

39. PLEIJ CWA, ABRAHAMS JP, VAN BELKUM A, PIETVELD K, BOSCH L: **The Spatial Folding of the 3′ Noncoding Region of Aminoacylatable Plant Viral RNAs.** In *Positive Strand RNA Viruses* edited by Brinton MA, Rueckert RR [book]. New York: Alan R Liss Inc., 1987, pp 299–316.

40. •• GALLIE DR, WALBOT V: **RNA Pseudoknot Domain of Tobacco Mosaic Virus can Functionally Substitute for a Poly(A) Tail in Plant and Animal Cells.** *Genes Dev* 1990, 4:1149–1157.
An extremely interesting dissection of the functions of various domains in the 3′-NTR of a plant virus.

41. • PETTY ITD, FRENCH R, JONES RW, JACKSON AO: **Identification of Barley Stripe Mosaic Virus Genes Involved in Viral RNA Replication and Systemic Movement.** *EMBO J* 1990, 9:3453–3457.
An excellent synthesis of all that is known about BSMV, the functions of the various genes, and the heterogeneity of RNA 3 within a population of genomes.

42. JACKSON AO, HUNTER BG, GUSTAFSON GD: **Hordeivirus Relationships and Genome Organization.** *Annu Rev Phytopathol* 1989, 27:95–121.

43. SIMONS JF, HELLMAN U, PETTERSSON RF: **Uukuniemi Virus Segment S: Ambisense Coding Strategy, Packaging of Complementary Strands into Virions and Homology to Members of the Genus Phlebovirus.** *J Virol* 1990, 64:247–255.

44. •• DE HAAN P, WAGEMAKERS L, PETERS D, GOLDBACH R: **The S RNA Segment of Tomato Spotted Wilt Virus has an Ambisense Character.** *J Gen Virol* 1990, 71:1001–1007.
Cogent arguments for the inclusion of tomato spotted wilt virus with the family *Bunyaviridae*.

45. • KAKUTANI T, HAYANO Y, HAYASHI T, MINOBE Y: **Ambisense Segment 4 of Rice Stripe Virus: Possible Evolutionary Relationship with Phleboviruses and Uukuviruses (*Bunyaviridae*).** *J Gen Virol* 1990, 71:1427–1432.
Sequence analysis to show that the terminal sequences are self-complementary and very similar to those of phleboviruses and uukuviruses.

46. •• KAKUTANI T, HAYANO Y, HAYASHI T, MINOBE Y: **Ambisense Segment 3 of Rice Stripe Virus: the First Instance of a Virus Containing Two Ambisense Segments.** *J Gen Virol* 1991, 72:465–468.
First description of a virus with two ambisense segments.

EG Strauss and JH Strauss, Division of Biology, California Institute of Technology, Pasadena, California 91125, USA.