CAN PLANT RNA VIRUSES EXCHANGE GENETIC MATERIAL?

Jean-Christophe Boyer, Marie-Dominique Morch and Anne-Lise Haenni

Institut Jacques Monod, 2 Place Jussieu - Tour 43, 75251 Paris Cedex 05, France

1. INTRODUCTION

The problem of the exchange of genetic material among viruses is intimately related to that of the evolution of these viruses, as are also the acquisition or the loss of genetic material. Similarly, the concept of virus evolution is associated with that of the origin of viruses.

The origin of viruses and their evolution have become subjects of intense research and discussion, as evidenced by the recent explosion of review articles on these topics (Holland *et al.*, 1982; Reanney, 1982; Goldbach, 1986; Gibbs, 1987; King *et al.*, 1987; Taylor and Hershey, 1987; Goldbach and Wellink, 1988; Hodgman and Zimmern, 1988; Strauss and Strauss, 1988). We have therefore preferred to restrict the subject of the present paper to one phenomenon participating in viral evolution, namely recombination; moreover, because of our current interest in plant viruses, this survey deals essentially with recombination among plant RNA viruses of (+) polarity, with only passing reference to reports of recombination among animal RNA viruses that do not replicate via DNA provirus copies.

Historically, the concept of recombination in RNA was long obliterated by the observations 1) that the enzymes involved in DNA recombination have no effect on RNA, 2) that early reports of recombination in influenza virus were in fact due to rearrangements of genome segments (Hirst, 1962), and 3) that with the exception of picornaviruses (Hirst, 1962; Pringle, 1965; McCahon and Slade, 1981; Kirkegaard and Baltimore, 1986; reviewed in King *et al.*, 1987) and the naturally-occurring MDV RNAs originating from bacteriophage Qβ (reviewed in Van Duin, 1988) no RNA recombination between viral RNAs could be detected, even between the well-studied RNA phages (Horiuchi, 1975). Furthermore, among plant viruses, neither recombination in mixedly infected plants, nor defective interfering (DI) RNAs had been demonstrated until recently.

This paper first surveys the bases of viral evolution. It then reviews the examples of plant RNA viruses for which recombination is either supported by direct experimental evidence or strongly suggested by sequence analyses. Finally, it discusses the possible mechanisms involved in recombination. This is a theme of ever-growing interest, as homologies between plant and animal viruses become increasingly apparent with the accumulation of viral genome sequence data.

M. Grunberg-Manago et al. (eds.), *Evolutionary Tinkering in Gene Expression*

2. BRIEF SURVEY OF THE BASES OF VIRAL EVOLUTION

Although it was long considered that RNA virus recombination must be an extremely rare event (if it existed at all), it was clear that RNA viruses undergo mutations at a very high rate. This has been thought to reflect the error level of the RNA polymerase, since no mechanism of proof-reading is detected in RNA virus replication.

2.1. Point Mutations

The effect of single base changes stems from formal genetic and serial passaging studies on RNA phages, later followed by work on eukaryotic RNA viruses. It is important to bear in mind the fact that exact mutation rates are difficult to establish 1) since they may be counter-balanced by reversions favored by selection pressure, 2) since rates must be evaluated for a single base change that allows the mutant to grow to a reasonable final titer in the absence of competition with the wild-type parent virus, and 3) since site-dependent variations for mutations exist. In spite of these reservations, it is clear that mutation rates for RNA are several orders of magnitude greater (10^{-4}) than for DNA (10^{-6} in prokaryotes). This high mutation rate may explain the small size of the genome of RNA viruses (typically 3×10^3 to 2×10^4 nucleotides), since the larger the genome, the smaller the number of mutations tolerated. This may also explain genome segmentation observed among many viruses: it allows for small genome parts, and thus enables non-viable mutated genome parts to be discarded without elimination of the total genome (Reanney, 1982).

It can therefore be expected that all progeny RNA molecules of 10^4 bases statistically contain one point mutation. However, a large percentage of these may be phenotypically silent mutations if for instance, they affect the third position of codons. The situation is not as simple, because the base in the third position of the codons of viruses is not random for those amino acids corresponding to 2 or 4 codons (Grantham *et al.*, 1981), suggesting that other factors intervene to maintain specific codons; it becomes even more complex in the case of viral genomes using overlapping reading frames as a strategy of expression.

In human influenza virus, the gradual accumulation of point mutations leading to single amino acid sequence changes ("antigenic drift") appears to form the basis for the alteration of antigenic sites, such that they are no longer recognized by the immune system (reviewed in Air and Laver, 1986).

There is nonetheless strong selective pressure to maintain a relatively uniform population of viral RNA molecules and to discard non viable mutants, possibly by direct competition during virus growth. An interesting aspect of selection pressure is DI-modulated virus infection. DI particles are virus deletion mutants that interfere specifically with the replication and packaging of the parental (helper) virus (reviewed in Holland *et al.*, 1980; Perrault, 1981). It has been suggested that intracellular coevolution could occur between the genetic material of DI RNAs and that of their parental virus. From this point of view vesicular stomatitis virus (VSV) and its associated DI particles have certainly been the most thoroughly studied. Cave *et al.* (1985) noticed that if mice are infected with a mixture of standard infectious VSV and its DI particles, continuous dynamic interaction occurs between the two populations for several days and with a cyclic pattern. It appears that such coevolution proceeds indefinitly through multiple cycles of selection of virus mutants (referred to as Sdi) resistant to a certain DI particle, followed by mutants resistant to a newly predominant DI particle etc. In recent studies, DePolo *et al.* (1987) examined the molecular consequences of this noteworthy phenomenon by sequencing parts

of certain Sdi genomes. It was demonstrated that the polymerase-associated (NS) genes are not usually mutated. On the other hand, the nucleocapsid (N) genes slowly accumulate stable mutations that could sometimes lead to large variations in the Sdi phenotype. Therefore, it seems likely that such recurring RNA population interactions can increase the rate of modifications of viral genomes and thus play an important role in virus evolution.

2.2. Recombination

This is the second important mechanism participating in virus evolution. Recombination basically concerns the formation of new combinations of genes or of linked genes resulting in new heritable characters or new combinations of such characters. Geneticists and molecular biologists often extend this definition to include the following phenomena.

1. Reassortment of genomic RNAs
2. Recombination of functional domains
3. Acquisition of exogenous genetic material.

These three basic types of "recombination" are discussed here; as will become apparent from the examples provided, they are not mutually exclusive.

2.2.1. Reassortment of Genomic RNAs

This corresponds to genetic rearrangement and is often referred to as "antigenic shift". It only occurs among viruses with a segmented genome, where "new" viruses can be obtained *in vivo* as well as in the test tube that possess new combinations of characteristics due to the exchange of genome parts (reviewed in Bruening, 1977; Air and Laver, 1986; Van Vloten-Doting and Jaspars, 1977; Mayo, 1987; Hiruki, 1987); these new isolates are called "pseudorecombinants". As demonstrated by back-cross experiments, the genome parts of the pseudorecombinants replicate true to type. In influenza virus, reshuffling among genome parts of two strains seems to be responsible for the appearance of "new" viruses infecting birds and man (reviewed in Webster *et al.*, 1982). The new isolates are not recombinants in the original sense of the term, since no breaking and rejoining of nucleic acid strand are considered to have occurred; for this reason, they will not be discussed further.

2.2.2. Recombination of Functional Domains

One of the main results of sequence comparisons among RNA viruses is the definition of protein domains conserved to various extents throughout several virus families and apparently characteristic of a given function. This section briefly describes the most significant domains and considers their combinations in various examples of viral genome since all these observations are in support of the "modular theory" of evolution which will then be presented.

2.2.2a. Functional Domains do Exist. A conserved domain is usually defined by a core consensus amino acid (aa) sequence thought to be required for its function. The aa sequence around this consensus shows additional homologies and specific residues that are conserved only in a limited number of viruses. These second order consenses have contributed to the grouping of virus families from animal and plant hosts into several "supergroups" or "superfamilies" (reviewed in Goldbach, 1986; Goldbach and Wellink, 1988; Hodgman and Zimmern, 1988; Strauss and Strauss, 1988). To date, the most spectacular information has been

deduced from comparative studies performed with (+) RNA viruses and only these will be discussed here.

Two supergroups of (+) RNA viruses among those not requiring a DNA step are now well defined on the bases of certain homologous domains and genome organization.

The "picorna-like" viruses (reviewed in Goldbach and Wellink, 1988) share with authentic picornaviruses (polio-, encephalomyocarditis-, rhinoviruses) the following features. A small viral protein (VPg) is bound at the 5' terminus of their genomic RNAs which are polyadenylated at the 3' end. Each genomic RNA is translated as a polyprotein subsequently cleaved to yield structural and non structural viral proteins. Among plant viruses, the como- (e.g. cowpea mosaic virus: CPMV), nepo- and possibly potyviruses are members of the picorna-like supergroup.

The "Sindbis-like" supergroup (reviewed in Goldbach and Wellink, 1988) associates the animal alphaviruses (Sindbis virus, Semliki Forest virus: SFV), possibly the coronaviruses (Gorbalenya *et al.*, 1988b) and a large variety of plant RNA viruses such as alfalfa mosaic virus (AlMV), the bromoviruses (e.g. brome mosaic virus: BMV), the cucumoviruses (e.g. cucumber mosaic virus: CMV), the furoviruses (e.g. beet necrotic yellow vein virus: BNYVV), the hordeiviruses (e.g. barley stripe mosaic virus: BSMV), the ilarviruses (e.g. tobacco streak virus, TSV), the potexviruses, the tobamoviruses (e.g. tobacco mosaic virus: TMV), the tobraviruses (e.g. tobacco rattle virus: TRV), the tymoviruses (e.g. turnip yellow mosaic virus: TYMV; Morch *et al.*, 1988) and possibly the carmoviruses (e.g. carnation mottle virus: CarMV) and the tombusviruses (e.g. cucumber necrosis virus: CNV). The genomes all have a cap structure at their 5' end. They all produce subgenomic RNAs in combination with various additional translation strategies to express their proteins (Morch *et al.*, 1987).

The first domain considered here has been found in all (+) RNA viruses. It is called the "polymerase" domain because of strong evidences in favor of a role of the corresponding proteins in nucleotide (nt) polymerization during viral replication. It is defined by a GDD (the one-letter code for amino acids is used when dealing with consensus sequences) consensus surrounded by hydrophobic residues (Kamer and Argos, 1984) and preceded 20 to 50 aa upstream by a conserved sequence [S/T]GxxxTxxxN[S/T]. Other conserved motifs exist that restrict the homology to one or the other virus supergroup (reviewed in Hodgman and Zimmern, 1988).

A second domain is also well characterized and very ubiquitous. It is designated the nucleoside triphosphate (NTP)-binding site, based on direct and indirect evidences that the corresponding proteins are able to bind NTP (reviewed in Gorbalenya *et al.*, 1988b). The consensus core of the domain is the sequence GxxGxGK[S/T]. It is followed at a distance that can vary from one virus group to the other by a degenerated consensus constituted of a D residue preceded by 3 hydrophobic residues. These two consenses together form part of larger domains spanning over 250 aa and 130 aa in the Sindbis-like and picorna-like supergroups respectively. These domains can be aligned throughout most members of each supergroup and reveal as many as 21 aa in the former and 45 aa in the latter supergroup, that are conserved, i.e. either invariant or having similar biochemical properties (Gorbalenya *et al.*, 1988b). Surprisingly such a domain is lacking in some viruses such as barley yellow dwarf virus (a luteovirus; Miller *et al.*, 1988), CarMV and CNV (Goldbach and Wellink, 1988). Thus, classification of these viruses in the Sindbis-like supergroup may need to be revised.

Other functional domains have been described in (+) RNA virus-coded proteins that are specific for a given supergroup or a given ensemble of virus families. One such domain is the proteinase domain found in picorna and picorna-like viruses. The proteolytic activity required by these

viruses for the processing of their polyprotein is encoded in the polyprotein itself; it has been characterized as a cysteine proteinase different from all cellular cysteine proteinases known to date. The essential C residue is present in a consensus sequence Txx[E/G]xCG[S/G]x[L/V/I] followed 5 to 10 aa downstream by a second consensus [I/V]xGxH where the H residue could be part of the C- and H-containing active site of the proteinase (Argos *et al.*, 1984; Gorbalenya *et al.*, 1986; Domier *et al.*, 1987).

Finally a domain has recently been characterized as specific for enzymes involved in DNA unwinding and may be called the helicase domain. It is present in all members of the Sindbis-like supergroup where it overlaps the NTP binding domain (Gorbalenya *et al.*, 1988a; Hodgman, 1988).

A completely different kind of conserved domain is found in the capsid proteins of (+) RNA viruses possessing isometric particles. All the coat proteins involved in such particles and studied so far contain a similar structural motif described as an eight-stranded anti-parallel β-barrel, referred to here as the β-barrel domain (reviewed in Gibbs, 1987). This motif is not found in the coat protein of TMV which forms helicoïdal particles. The definition of a conserved structural motif for the coat proteins of filamentous or rod-shaped RNA viruses awaits further studies on these proteins.

The clearly established existence of conserved domains leads to several speculations. It could for example favor the hypothesis of divergent evolution according to which all protein sequences associated with a given function derive from a common ancestral sequence. The purpose here is rather to use the various domains as units of genetic information for viral genomes and analyze their recombination potential.

2.2.2b. Different Organizations of Functional Domains. The relative association of one domain with regard to the other can vary according to three parameters, 1) the nature and the number of the domains that are associated, 2) their organization in the genome, and 3) their mode of expression. Examples of these various combinations follow.

In the Sindbis-like supergroup, the same polymerase domain can be found in isometric particles (AlMV, BMV, CMV, Sindbis virus) as well as in helicoïdal particles (TMV). This domain is thus combined in the former case with the β-barrel domain specific for isometric particles and in the latter case (helicoïdal particles) with a domain specifying another type of coat protein. Similar combinations (Gibbs, 1987) are possible among the picorna-like viruses which possess either isometric (picornaviruses, comoviruses), or rod-shaped (potyviruses) virion particles.

Reciprocally one should stress that the same β-barrel motif can be found associated with Sindbis-type as well as with picorna-type functional domains.

In alphaviruses a domain coding for the glycoproteins engaged in the lipid envelope of the virion is also added to the capsid (Fuller and Argos, 1987). Recently evidence has been provided that recombination has brought together the capsid domain of one alphavirus species with the glycoprotein domain of another species and thus has given rise to a third alphavirus species (Hahn *et al.*, 1988).

Interestingly, the conserved structural motif is present in one copy in the genome of the isometric viruses of the Sindbis-like supergroup and in three copies in the isometric viruses of the picorna-like viruses (Rossmann *et al.*, 1985). These three β-barrel motifs are carried either by three different proteins (VP1, VP2 and VP3) in picornaviruses or by two proteins only (VP37 and VP23) in CPMV (Goldbach, 1986).

Other domains can be present twice on a genome. Two proteinase domains are found in some of the picorna-like viruses (proteinases 2A and 3C of poliovirus, protein 32K and proteinase 24K of CPMV). One of the two copies may have lost its original function and thus have diverged from

the consensus (case of the 32K protein of CPMV; reviewed in Wellink and Van Kammen, 1988). Two NTP-binding domains are also found among some members of the Sindbis-like supergroup (BNYVV, BSMV). Here again one copy of the domain is less conserved than the other suggesting diversification through independent evolution following duplication (Goldbach and Wellink, 1988; Gorbalenya *et al.*, 1988a).

Differences in genome organization with respect to the polymerase and the NTP-binding domains appear clearly in the Sindbis-like supergroup. In tripartite genomes (AlMV, BMV, CMV, TSV), these two domains are carried by two separated RNAs, whereas in monopartite genomes such as TMV or Sindbis virus, they are carried by the same RNA.

In TMV, the two domains remain associated on a single large polypeptide whereas in Sindbis virus they become dissociated by proteolytic cleavage. Moreover, sequence alignments reveal in Sindbis virus the insertion of an additional as yet non-assigned domain between the polymerase and the NTP-binding domains which are contiguous in TMV (Ahlquist *et al.*, 1985).

Picornaviruses and CPMV also differ in the sense that the former have a monopartite and the latter a bipartite genome. However the relative organization of the functional domains remains very parallel between these two types of genome (Goldbach, 1986).

In monopartite genomes the capsid β-barrel domains can be found either at the 3' end (e.g. tymoviruses), at the 5' end (e.g. picornaviruses), or in internal regions of the genome (e.g. tombusviruses).

The mode of expression of a given domain may be different among the members of a same virus family. The polymerase domain of the Sindbis virus is a cleavage product of a polyprotein precursor that itself derives from the readthrough of a UGA termination codon located just upstream of the region coding for the polymerase domain. However in SFV, the polyprotein precursor for the polymerase domain is translated without interruption by any termination codon (reviewed for the other alphaviruses in Strauss *et al.*, 1987).

A similar situation, although between two different virus families, has recently been observed when comparing TMV and TYMV. Readthrough of a termination codon is involved in the former virus to account for the translation of the polymerase domain whereas continuous translation of an uninterrupted open reading frame (ORF) in the latter case yields a polyprotein containing the polymerase domain (Morch *et al.*, 1988).

2.2.2c. The Hypothesis of "Modular Evolution". Functional domains have been defined and are associated in various combinations throughout many virus families. This statement recalls the hypothesis of modular evolution proposed by Botstein (1980) on the basis of studies performed with the DNA bacteriophages λ and P22. As reported by Gibbs (1987), this hypothesis is defined by the "joint evolution of sets of functionally and genetically interchangeable elements (the domains or modules), each of which carries out a particular biological function". In other words (Hodgman and Zimmern, 1988), "the distinct virus groups are related by the association of shared common components with unique components (such as the transport proteins in plant viruses) that define the specific features of this group". In the case of retroviruses, the possibility is not excluded that rearrangements of modules have taken place during the DNA step of the viral cycle. However for the other RNA viruses neither DNA intermediates nor pseudogenes of viral genomes have been described. Thus, whatever the origin and evolution of a given virus family, the fact remains that the interpretation of all sequence analyses performed on contemporary viral proteins strongly suggests that recombination and gene duplication may occur at the RNA level in viral genomes.

2.2.3. Acquisition of Exogenous Genetic Material

Recombination has been considered so far between RNA genomes but exchanges may also occur between cellular and viral genetic material. However not all cellular sequences are known and divergence has occurred since their putative integration into the viral genome. Thus it is often very difficult to trace such sequences in a viral genome. Indirect evidence derived from comparative sequence analyses and more direct evidences based on very limited observations attempt to support the statement of recombination between cellular and viral sequences.

2.2.3a. Homologies of Viral Domains with Cellular Proteins. The comparison of the functional domains described above with cellular proteins has stressed striking homologies. The NTP-binding domain of (+) RNA viruses especially in the Sindbis-like supergroup can be aligned with a series of cellular prokaryotic and eukaryotic proteins that have in common the property of interacting with nucleic acids and unwinding helical turns (Hodgman, 1988, Gorbalenya *et al.*, 1988a; Lane, 1988).

The proteinase of picorna-like viruses, although claimed to be a new type of cysteine proteinase can be aligned with both cysteine and serine proteinases suggesting it could derive from a common ancestor of both cellular proteinases (Gorbalenya *et al.*, 1986).

Other homologies of virus-specific sequences with cellular sequences may still be revealed. For example the basic N-terminal region of several but not all capsid proteins of isometric viruses has been proposed to be related to histone or protamine sequences (Argos, 1981).

Although the hypothesis of convergent evolution of sequences is not excluded, these reports are to be taken as indications for the possible cellular origin of some viral sequences. Some direct observations may give weight to this hypothesis.

2.2.3b. Direct Evidence of Transfer from Host to Virus. A very well known system of genetic recombination between host and virus is the transducing ability of retroviruses such as the viral copies of oncogenes (v-*onc*) which originated from cellular copies of the same genes (c-*onc*) that were integrated into a retroviral genome (reviewed in Bishop, 1983). This type of recombination is likely to occur at the DNA level and thus will not be further developed here.

An interesting case is however that of some DI particles of Sindbis virus (Monroe and Schlesinger, 1983). Sequence analyses of these DI genomes established the presence at their 5' end of part of a cellular $tRNA^{Asp}$ (see 3.6.). Another example is found in influenza virus: transcription of the mRNAs of this (-) RNA virus is primed with capped oligonucleotides derived from cellular mRNAs (Plotch *et al.*, 1981). In both cases the extent of cellular sequences combined with the viral-related sequences is very small. These examples provide indications that during the virus cycle viral RNAs come in contact with cellular RNAs.

Further speculations on the possible involvement of cellular RNA sequences in the constitution of viral particles relies on more knowledge about RNA recombination mechanisms (see 4.).The possible existence of "cellular RNA genetic elements" has been postulated (Zimmern, 1982) that would be capable of being replicated, as are for example viroids, and of being integrated into viral RNAs via mechanisms that could be the reversal of splicing mechanisms.

3. EXAMPLES OF RECOMBINATION AMONG PLANT RNA VIRUSES

This section reviews the examples of (likely) recombination among

plant viruses: the first is the only well-established example of recombination, whereas the other four correspond to very probable recombination events. The sixth part briefly reviews a few well-studied examples of recombination among animal RNA viruses and their DI RNAs.

3.1. RNA3 of Brome Mosaic Virus (BMV)

This is the only clear-cut example of exchange of genetic material among plant RNA viruses reported to date (Bujarski and Kaesberg, 1986; reviewed in Ahlquist *et al.*, 1987).

The genome of BMV is composed of three distinct RNA molecules (designated RNA1, 2 and 3 in the order of their decreasing size) encapsidated in separate virions. All three components are necessary to induce detectable systemic infection in whole barley plants. The sequence of each RNA has been established. The 3' terminal 193 nt of each RNA are 97-99% homologous with the other two components, suggesting that their conservation is under considerable selection pressure. This 3' region is responsible for the tRNA-like properties of the viral RNA and it also directs the synthesis *in vitro* of the (-) RNA strand by the virus-specific RNA polymerase complex.

Advantage has been taken of these conserved regions to probe RNA recombination in plant cells. An engineered BMV RNA3 (designated m4 RNA3) has been constructed in which nt 81-100 from the 3' end of RNA3 have been deleted. This region corresponds to a stem and loop that is normally present at the 3' end of the RNAs of BMV; it is conserved in the RNAs of cowpea chlorotic mottle virus, but absent from the RNAs of broad bean mottle virus, two other bromoviruses. When barley plants are inoculated with wild-type (wt) BMV RNA1 and 2 and the deleted m4 RNA3, accumulation of progeny RNA is lower than if m4 RNA3 is replaced by wt RNA3. Interestingly, in systemic infection, m4 RNA3 is unstable: multiple pseudorevertants appear in which the deleted sequence has been restored due to recombination between m4 RNA3 and either wt RNA1 or RNA2. This is accompanied by a concomittant decrease in m4 RNA3.

A close analysis of the pseudorevertants has revealed that the stem and loop region has been regained by RNA3. Each revertant also bears one or more base changes 3' of the restored sequence and a few additional changes 5' of this sequence. These changes correspond to sequences present in either RNA1 or RNA2, so that it is possible to ascribe the pseudorevertants as resulting from a recombination event between RNA3 sequences upstream of the missing stem and loop, and 3' sequences of RNA1 or RNA2. Of the five pseudorevertants analyzed at the molecular level, three result from recombination within the highly conserved region of 193 nt. The other two result from recombination outside of this homologous region, demonstrating that recombination is not limited to homologous sequences (see 4.2.).

3.2. RNA2 of Tobacco Rattle Virus (TRV)

TRV contains a bipartite genome, RNA1 and RNA2. The viruses of this group are divided into three clusters, cluster 1 (strains of TPV), cluster 2 (strains of PEBV) and cluster 3 (strain CAM of TRV). Whereas only the 3' terminal third of the sequence of RNA1 is known, RNA2 has been totally sequenced for representatives of all three clusters. The length of RNA2 differs from strain to strain: it is 1799 nt long for strain CAM, 1905 nt for strain PSG of cluster 1, and 3389 nt for strain TCM, also of cluster 1.

The 3' termini of CAM-RNA2, PSG-RNA2 and TCM-RNA2 show 100% homology with the corresponding RNA1 over a length of 459, 497 and 1099 nt respectively, indicating a strong selection pressure in this region and suggesting recombination between the 3' region of RNA1 and RNA2. The most

striking case is that of TCM-RNA2. Its 1099 nt-long RNA stretch homologous to RNA1 contains the C-terminal region of the 28.8 K (K = kilodalton) ORF of RNA1, followed by the 16 K ORF, the 3' proximal ORF of RNA1 (Cornelissen *et al.*, 1986; Angenent *et al.*, 1986).

3.3. tRNA-Like Regions in Tobamoviruses

Tobamoviruses contain a monopartite RNA genome whose 3' terminal region possesses tRNA-like properties. These viruses have been divided into two subgroups, on the basis of the location of the assembly origin (reviewed in Takamatsu *et al.*, 1983). The common (*Vulgare*) strain and the tomato strain (L strain) belong to subgroup 1 (assembly origin is 800-1000 nt from the 3' end), whereas the cowpea strain (Cc strain) and cucumber green mottle mosaic virus (CGMMV) belong to subgroup 2 (assembly origin 300-500 nt from the 3' end).

The complete sequence of the genome of the *Vulgare* and L strains, and the sequence of the 3' region of the genome of the Cc strain and CGMMV are known, allowing comparisons between the 3' non coding regions of these RNAs (reviewed in Morch and Haenni, 1987). Furthermore, the RNA of the *Vulgare* strain and of CGMMV accepts histidine, whereas that of the Cc strain accepts valine (reviewed in Joshi *et al.*, 1983b).This is an interesting situation because in all other cases of aminoacylation among plant RNA viruses, the RNA of the viruses belonging to a given taxonomic group accept the same aa. On the other hand, the RNA of tymoviruses such as TYMV accepts valine.

A search for sequence similarities has been made between the tRNA-like regions of the *Vulgare* and Cc strains, since they both belong to tobamoviruses but accept different aa, and between the tRNA-like regions of TYMV and the Cc strain, since they belong to different taxonomic groups but accept the same aa (Joshi *et al.*, 1983a). The overall tRNA-like region of TYMV and the Cc strain presents considerable sequence conservation, suggesting that among these conserved regions lie recognition sites for the Val-tRNA synthetase. In contrast, sequence similarity between the tRNA-like regions of the *Vulgare* and Cc strains is confined to the first 3' coterminal $\cong 42$ nt; this conserved region could be required for interaction with other tRNA-specific enzymes, and/or with the RNA polymerase.

The 3' terminal sequences of the CGMMV and L strains are extensively homologous to that of the *Vulgare* strain, consistant with the observation that the RNA of the *Vulgare* strain and of CGMMV accept histidine; it can thus be proposed with confidence that the 3' region of the L strain also accepts histidine, although this has not been tested.

The sequence similarities observed between the 3' region of the Cc strain and TYMV on one hand, and between the 3' region of the Cc strain and the other tobamoviruses on the other hand suggest that the Cc strain may have arisen by recombination between the coding body of a tobamovirus such as the *Vulgare* strain, and the tRNA-like non-coding 3' region of a tymovirus.

3.4. Satellite of Turnip Crinkle Virus (TCV)

Satellite RNAs are small molecules that require the company of a helper virus for their replication. Their presence in the plant can modify (intensify or reduce) the symptoms produced by the helper virus alone. Most satellites are encapsidated in helper virus particles, although in certain cases they are encapsidated in separate particles which they themselves encode. They vary in size (200-1700 nt), coding capacity, and form (linear or circular). There is little in common between satellites that accompany viruses of different groups. Moreover, there is little homology between satellites and their helper virus genomes, a feature that is said to distinguish satellites from DI

particles (reviewed in Murant and Mayo, 1982; Francki, 1985; Morch and Haenni, 1987).

However, an unexpected situation had recently been described for RNA C, a virulent satellite of TCV (Simon and Howell, 1986). Its 355 nt-long linear sequence is composed of 2 major domains of about equal lengths. The 3' domain presents 93% homology with the 3' non-coding region of TCV RNA, whereas the 5' domain is homologous to other satellites of TCV. Thus RNA C appears to be a composite molecule with elements of the TCV genome and of other TCV satellites. It thus shares features of DI RNAs and conventional satellites, and it seems possible that it may have arisen from recombination between TCV RNA and one of the other satellites of this virus.

3.5. DI RNA of Tomato Bushy Stunt Virus (TBSV)

To date, only one report has clearly established the existence of a DI particle in association with a plant virus (Hillman *et al.*, 1987). Preparations of TBSV contain a 0.4 kilobase-long, genome-related and symptom-modulating species whose characteristics comply with those of DI RNAs. This DI RNA is a colinear deletion mutant of the parental virus and it is a mosaic of 7 juxtaposed RNA stretches ranging in size from 5 to 130 nt in the same relative order as in the parental virus.

3.6. Brief Survey Among Animal RNA Viruses and Their DI RNAs

Among picornaviruses, recombination was first described in poliovirus (Hirst, 1962; Ledinko, 1963) and soon thereafter (Pringle, 1965) in foot-and-mouth disease virus (FMDV). More recently evidence for recombination has been obtained for mouse hepatitis virus (MHV), a coronavirus (Lai *et al.*, 1985).

Recombinants of the RNA genomes of poliovirus (Kirkegaard and Baltimore, 1986), FMDV (King *et al.*, 1982; reviewed in King *et al.*, 1987) and MHV (Keck *et al.*, 1988)have been analyzed in detail by T1 fingerprinting and/or nt sequencing.

A recent example of possible natural recombination has been reported (Hahn *et al.*, 1988) with respect to an alphavirus species that contains domains of two other alphavirus species (see *2.2.2b.*).

Among animal viruses, DI particles have been found associated not only with minus-stranded RNA viruses such as VSV, influenza virus and Sendai virus (reviewed in Lazzarini *et al.*, 1981), but also with positive-stranded RNA viruses such as the alphaviruses Sindbis virus and SFV (Stollar, 1980) and poliovirus (Kuge *et al.*, 1986).

The structure of these DI RNAs generally involves complex recombinations of parental viral sequences (Lazzarini *et al.*, 1981; Söderlund *et al.*,1981; Tsiang *et al.*, 1988). This can best be illustrated by the mosaic DI RNA of influenza virus: it is composed of one region of segment 1 flanked by different regions of segment 3. In addition, the regions of segment 3 are borrowed alternatively from the 5' and the 3' part of that segment (Fields and Winter, 1982).

DI RNA formation can sometimes encompass the acquisition of non viral sequences. From this point of view, the example of several DI RNAs isolated from Sindbis virus-infected cells is striking: studies of these DI RNAs have revealed that they possess at their 5' termini the 67 nt deriving from the 3' end of a cellular $tRNA^{Asp}$ covalently attached to either nt 23 or 31 of the virion RNA (Monroe and Schlesinger, 1983, 1984). The exact significance of this surprising feature remains unclear since it does not seem to provide any selective advantage to this DI RNA population over those lacking the host sequence (Tsiang *et al.*, 1985, 1988).

4. POSSIBLE MECHANISMS OF RECOMBINATION

The exact mechanisms of recombination are still the subject of speculation. However, two major mechanisms can be distinguished, splicing and the "copy-choice" mechanism.

4.1. Splicing and the Copy-Choice Mechanism

Splicing, as postulated at least for the formation of some plant satellite RNAs, requires the pre-existence of a precursor RNA in which excision and ligation at consensus sequences occurs (reviewed in Symons *et al.*, 1985). Trans-splicing (Solnick, 1985; Konarska *et al.*, 1985) could also account for recombination between viral genomes or between a viral genome and a cellular RNA (Zimmern, 1982). To date among viral genomes and DI RNAs, no consensus sequences for splicing events have been observed at the junction of the recombinant regions. For this reason, the copy-choice mechanism is the favored model for recombination.

In the copy-choice mechanism, the polymerase stops synthesis on its template strand and pursues synthesis by reinitiating elsewhere without releasing the nascent daughter strand (Lazzarini *et al.*, 1981; Holland *et al.*, 1982; Kolakofsky and Roux, 1987; Emerson and Schubert, 1987). The polymerase can reinitiate synthesis further on the same strand such as by slippage of the enzyme along the template, or it can switch template and reinitiate synthesis on another copy of the same RNA molecule or on a different RNA molecule (Lazzarini *et al.*, 1981). The different RNA molecules can be anything: the nascent daughter strand itself leading for instance to "snap-back" particles (Lazzarini *et al.*, 1981), another segment in viruses with a multipartite genome as might occur in the case of RNA3 of BMV (see 3.1.) and of RNA2 of TRV (see 3.2.), the RNA of another coinfecting virus as for the Cc strain in the tobamoviruses (see 3.3.), the RNA of the helper virus in the case of the satellite of TCV (see 3.4.), and finally a cellular RNA such as the $tRNA^{Asp}$ of certain DI RNAs of Sindbis virus (see 3.6.). More than one reinitiation event is required in some cases to account for the formation of a recombinant RNA, such as the DI RNA of TBSV (see 3.5.). In the case of the DI RNA of influenza virus (see 3.6.), it has been proposed that the enzyme could zigzag between the 5' and 3' ends of one or two copies of segment(s) 3, assuming these regions are close to one another; in addition, intersegment recombination with segment 1 would occur to complete formation of the DI RNA. In MHV-infected cells, discrete RNA species containing the MHV leader sequences as well as larger leader-containing RNA species are found (discussed in Makino *et al.*, 1986a, 1986b). They might correspond to normal nascent intermediates of RNA replication, or to incomplete transcription products. The presence of such free RNA intermediates supports the copy-choice model of recombination.

4.2. Homologous Versus Non-Homologous Recombination

In an effort to define at the molecular level the elements that dictate recombination, attempts have been made to classify recombination events into homologous and non-homologous (reviewed in King *et al.*, 1987; Hodgman and Zimmern, 1988).

Homologous recombination occurs in regions of homology of the protagonist RNAs. As a result, the region of homology is present in the progeny RNA, making the exact site of recombination difficult to assess. A general alignment of the parental molecules with a certain degree of base-pairing, or the stabilization of a structure on one and/or the other side of the immediate cross-over point may favor homologous recombination. This type of recombination appears to be responsible for at least some recombination events observed among picornaviruses (Kirkegaard and Baltimore, 1986; reviewed in King *et al.*, 1987). Indeed,

the only region of the genome where recombination has not been observed in FMDV is the region with the most poorly conserved sequence: sequence matches of 18 and 32 bases are found in two cross-over points (King *et al.*, 1986). Homologous recombination is also postulated to be responsible for the generation of certain poliovirus-associated DI RNAs (Kuge *et al.*, 1986).

However, since sequence homology is not a prerequisite for all recombination events, non-homologous recombination must also be invoked. The essential difference between homologous and non-homologous recombination rests in the minimum length of base-pairing required for a primer chain to be elongated, and it has been suggested that this may depend on the polymerase (King *et al.*, 1987).

Non-homologous recombination is less well documented and has not been extensively studied. DI RNAs presumably result from this kind of recombination: few (if any) matching bases are detected at the cross-over point, the most common situation being one base match (Jennings *et al.*, 1983), with a preference for a purine residue at the 3' end of the homologous region. The DI RNA of TBSV (see 3.5.) and the mosaic DI of influenza virus (see 3.6.) are likely examples of such non-homologous recombinations among plant and animal viruses respectively. Another example is found among certain DI RNAs of Sindbis virus (see 3.6.): only a single matching base can be distinguished at the cross-over point between the $tRNA^{Asp}$ and the DI RNA sequence. Non-homologous recombination could account for the situation observed in the tRNA-like region of the Cc strain of TMV (see 3.3.), and the satellite of TCV (see 3.4.). However, it is also possible that a region of sequence homology existed previously in the two parent genomes that permitted homologous recombination to occur; with time, point mutations and evolutionary pressure would have obliterated such a sequence in at least one parent so as to no longer be detectable. This example points to the caution that is warranted in designing models of recombination based on sequence homologies.

Whatever the mode of recombination, the copy-choice model postulated for the generation of DI RNAs supposes that re-initiation of synthesis can be directed efficiently by a single base-pair formed between a 3'-terminal nt (preferentially an A or a G) in the primer and a complementary base in the template.

It is worth returning to the situation encountered in the RNA3 or BMV (see 3.1.). Of the five recombinants analyzed, three have occurred within the 193 nt of the conserved region, whereas the other two occurred upstream, in a region devoid of sequence homology. Thus the RNA polymerase of BMV has presumably performed both homologous and non-homologous recombinations: the mode of recombination does not seem to be an inherent property of at least this polymerase, as has been proposed for others (King *et al.*, 1987).

4.3. Dissecting Recombination

Two questions directly concern recombination: 1) the reasons for which the nascent polymerase-RNA complex leaves its original template, and 2) how it chooses its new template.

Pausing during elongation could cause the nascent polymerase-RNA complex to leave its template. This might occur because of a specific sequence in the template. Indeed, in the case of the mosaic DI RNA of influenza virus, it has been proposed (Fields and Winter, 1982) that uridine-rich regions in the template cause the polymerase to pause: if the polymerase-RNA complex were then to leave its original template, it might resume synthesis at another site. Secondary and tertiary structures within the template might also lead to pausing during elongation (reviewed in Lazzarini *et al.*, 1981). In addition, one cannot exclude the possibility that other factors, such as proteins (capsid proteins ?) might interfere with RNA elongation.

Whatever the reason for which the nascent polymerase-RNA complex is detached from its original template, in homologous recombination it is fairly easy to visualize how the complex might base-pair to a new template, allowing RNA synthesis to proceed.

In non-homologous recombination, the parameters dictating resumption of synthesis are far more elusive. Since base-pairing is not a prerequisite in this form of recombination, other elements are most likely involved. A simple model that can be proposed is one in which the nascent polymerase-RNA complex would bind to the new template at a site for which the polymerase would have a high affinity. In this respect, it is interesting that in strains PSG and TCM of TRV (see 3.2.), just downstream of the junction point between RNA1 and RNA2, the sequence AUAAUUGUU occurs; this sequence resembles the 5'-termini of the genomic and subgenomic TRV RNAs and could be the site of internal initiation for the replicase on minus-strand RNA (Angenent *et al.*, 1986). It would thus appear logical for the polymerase to bind to this sequence. An additional possibility is that structure rather than sequence could constitute the target for binding of the polymerase, as proposed in the case of the DI RNAs of Sindbis virus (Tsiang *et al.*, 1988); indeed, the 5'-terminal region in these DI RNAs can undergo extensive variations while still maintaining the ability to be replicated by the viral-coded enzyme.

Pausing and reinitiation of synthesis at a new template site might be brought about by way of certain transient highly ordered structures that would bring potential cross-over regions in close proximity with one another. This could be brought about by looping-out of sequences that would consequently be skipped by the polymerase. A computer-aided search of poliovirus DI RNAs has revealed a secondary structure present in every deletion point in the parental genome that could be involved in DI RNA formation (Kuge *et al.*, 1986) In this model, the region(s) at which pausing and re-start would occur (either on the same or on a different RNA template) would be maintained together by "supporting" RNA sequences. Such supporting sequences could reside within poliovirus-specific RNA sequences themselves. This "supporting sequence-loop" model suggests that specific interactions exist between interruption and resumption sites. The model has been developed for poliovirus-associated DI RNAs, but it might be extended to explain the formation of DI RNAs of other viruses. Indeed, a similar computer-aided search has been made of rearranged sequences of DI RNAs of influenza virus and of Sindbis virus: supporting sequences for every rearranged site have been discovered in these mutated genomes (Kuge *et al.*, 1986).

5. CONCLUSIONS

One sould bear in mind that the recombinant RNAs we observe most likely only represent a fraction of those that are produced by Nature, since only the viable RNA progeny that has acquired a sufficient level of replication capacity and encapsidation is detectable.

For recombination to occur in Nature between distinct parental viral RNAs, it is obviously mandatory for the same host cell to allow both parents to replicate. This is probably not a difficult constraint for related viruses that in any event develop in the same host. On the other hand, this condition becomes more difficult to meet when the viruses are unrelated. One can postulate that recombination has occurred between the genomes of two plant viruses belonging to different families in the case of the Cc strain of TMV (see 3.3.), if indeed the genome of this virus results from recombination between a tobamovirus and a tymovirus. An investigation of the propagation hosts of tymoviruses suggests that this is not unlikely, since several tymoviruses replicate in tobacco plants or in peas (Blok *et al.*, 1987) that are also hosts of tobamoviruses.

A further level of complexity arises if one considers recombination between plant and animal viruses, since such an event also requires that both viruses replicate in the same host cell. A feature that is common to several virus families of both kingdoms is their propagation by insect vectors. It has been proposed (Goldbach, 1986) that insects may have harboured common ancestor viruses from which the present-day viruses derive. Indeed, the insect host range of plant and animal viruses overlaps, and certain insect RNA viruses also infect mammalian cells, such as black beetle virus (a nodamura virus), an insect virus with a divided genome, a common feature among plant viruses.

To complete the picture of the viruses discussed here, the demonstration of a plant RNA virus (single-stranded and of positive polarity) also capable of replicating in an insect or vice versa, is still lacking. A virus endowed with this capacity may be difficult to observe if the pressure of evolution were to have modified it to such an extent that it no longer replicates in one or the other host. Such a possibility is supported by studies with wound tumor virus: if this plant virus is propagated in its host plant for up to two years without passage through an insect vector, mutants appear that are no longer capable of being transmitted by the insect. Such mutants lack segments of their genome: the host range has thus been changed by changing the ecology of the virus (reviewed in Taylor and Hershey, 1987).

What advantage does recombination represent for RNA viruses? It is likely that recombination is functionally important, particularly for viruses with unsegmented genomes. No proof-reading mechanism exists for RNA genomes that might correct errors made by the RNA polymerase. However, "correction" by recombination rather than by proof-reading might circumvent this problem and could help to establish a population consensus. Recombination might also constitute a natural source of variants of a given virus, and from this point of view could play a similar role in virus evolution as reassortment of genome parts plays for viruses with a segmented genome. Finally, by enabling the virus to draw on host material, recombination offers a new dimension for viral variation.

ACKNOWLEDGMENTS We thank F. Chapeville for his constant interest and his encouragements. J.C.B. is grateful to the "Ministère de la Recherche et de l'Enseignement Supérieur" for a fellowship. This work was financed in part by a grant from the "A.I.P. N° 87/4560 - Obtention Plantes Résistantes aux Virus". The Institut Jacques Monod is an "Institut Mixte, C.N.R.S. - Université Paris VII".

REFERENCES

Ahlquist, P., French, R., and Bujarski, J. J., 1987, Molecular studies of brome mosaic virus using infectious transcripts from cloned cDNA, *Adv. Virus Res.* **32**:215-242.

Ahlquist, P., Strauss, E. G., Rice, C. M., Strauss, J. H., Haseloff, J., and Zimmern, D., 1985, Sindbis virus proteins nsP1 and nsP2 contain homology to nonstructural proteins from several RNA plant viruses, *J. Virol.* **53**:536-542.

Air, G. M., and Laver, W. G., 1986, The molecular basis of antigenic variation in influenza virus, *Adv.Virus Res.* **31**:53-102.

Angenent, G. C., Linthorst, H. J. M., Van Belkum, A. F., Cornelissen, B. J. C., and Bol, J. F., 1986, RNA 2 of tobacco rattle virus strain TCM encodes an unexpected gene, *Nucl. Acids Res.* **14**:4673-4682.

Argos, P., 1981, Secondary structure prediction of plant virus coat proteins, *Virology* **110**:55-62.

Argos, P., Kamer, G., Nicklin, M. J. H., and Wimmer, E., 1984, Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families, *Nucl. Acids Res.* **12**:7251-7267.

Bishop, J.M., 1983, Cellular oncogenes and retroviruses, *Ann. Rev. Biochem.* **52**:301-354.

Blok, J., Gibbs, A., and Mackenzie, A., 1987, The classification of tymoviruses by cDNA-RNA hybridization and other measures of relatedness, *Arch. Virol.* **96**:225-240.

Botstein, D., 1980, A theory of modular evolution for bacteriophages, *Ann. N. Y. Acad. Sci.* **354**:484-491.

Bruening, G., 1977, Plant covirus systems: two-component systems, in: *"Comprehensive Virology"*, H. Fraenkel-Conrat, and R. R. Wagner, eds., Plenum Press, New York, vol. **11**, pp. 55-141.

Bujarski, J. J., and Kaesberg, P., 1986, Genetic recombination between RNA components of a multipartite plant virus, *Nature* **321**:528-531.

Cave, D. R., Hendrickson, F. M., and Huang, A. S., 1985, Defective interfering virus particles modulate virulence, *J.Virol.* **55**:366-373.

Cornelissen, B. J. C., Linthorst, H., J., M., Brederode, F., Th., and Bol, J. F., 1986, Analysis of the genome structure of tobacco rattle virus strain PSG, *Nucl. Acids Res.* **14**:2157-2169.

DePolo, N. J., Giachetti, C., and Holland, J. J., 1987, Continuing coevolution of virus and defective interfering particles and of viral genome sequences during undiluted passages:virus mutants exhibiting nearly complete resistance to formely dominant defective interfering particles, *J.Virol.* **61**:454-464.

Domier, L. L., Shaw, J. G., and Rhoads, R. E., 1987, Potyviral proteins share amino acid sequence homology with picorna-, como- and caulimoviral proteins, *Virology* **158**:20-27.

Emerson, S. U.and Schubert, M., 1987, Molecular basis of rhabdovirus replication, in *"The Molecular Basis of Viral Replication"*, R. Perez Bercoff, ed., Plenum Press, New York, London, pp. 255-276.

Fields, S., and Winter, G., 1982, Nucleotide sequences of influenza virus segments 1 and 3 reveal mosaic structure of a small viral RNA segment, *Cell* **28**:303-313.

Francki, R. I. B., 1985, Plant virus satellites, *Ann. Rev. Microbiol.* **39**:151-174.

Fuller, S. D., and Argos, P., 1987, Is Sindbis a simple picornavirus with an envelope?, *EMBO J.* **6**:1099-1105.

Gibbs, A., 1987, Molecular evolution of viruses; 'trees', 'clocks' and 'modules', *J. Cell. Sci. Suppl.* **7**:319-337.

Goldbach, R. W., 1986, Molecular evolution of plant RNA viruses, *Ann. Rev. Phytopathol.* **24**:289-310.

Goldbach, R., and Wellink, J., 1988, Evolution of plus-strand RNA viruses, *Intervirology*, submitted for publication.

Gorbalenya, A. E., Blinov, V. M., and Donchenko, A.P.,1986, Poliovirus-encoded proteinase 3C: a possible evolutionary link between cellular serine and cysteine proteinase families, *FEBS Lett.* **194**:253-257.

Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P., and Blinov, V. M., 1988a, 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.

Gorbalenya, A. E., Blinov, V. M., Donchenko, A.P., and Koonin, E.V., 1988b, 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.

Grantham, R., Gautier, C., Gouy, M., Jacobzone, M., and Mercier, R., 1981, Codon catalog usage is a genome strategy modulated for gene expressivity, *Nucl. Acids Res.* **9**:r43-r74.

Hahn, C. S., Lustig, S., Strauss, E. G., and Strauss, J. H., 1988, Western equine encephalitis virus is a recombinant virus, *Proc. Natl. Acad. Sci. USA*, **85**:5997-6001.

Hillman, B. I., Carrington, J. C., and Morris, T. J., 1987, A defective interfering RNA that contains a mosaic of a plant virus genome, *Cell* **51**:427-433.

Hirst, G.K., 1962, Genetic recombination with Newcastle disease virus, polioviruses and influenza, *Cold Spring Harbor Symp. Quant. Biol.* **27**:303-309.

Hiruki, C., 1987, The dianthoviruses: a distinct group of isometric plant viruses with bipartite genome, *Adv. Virus Res.* **33**:257-300.

Hodgman, T. C., 1988, A new superfamily of replicative proteins, *Nature* **333**:578.

Hodgman, T. C., and Zimmern, D., 1988, Evolution of RNA viruses, in: *"RNA Genetics"*, J. Holland, E. Domingo, and P. Ahlquist, eds., CRC Press, Boca Raton, Fla., in press.

Holland, J. J., Kennedy, S. I. T., Semler, B. L., Jones, C. L., Roux, L. and Grabau, E. A., 1980, Defective interfering RNA viruses and host-cell response, in: *"Comprehensive Virology"*, H. Fraenkel-Conrat and R. R. Wagner, eds., Plenum Press, New York, vol. **16**, pp. 137-192.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol, S., 1982, Rapid evolution of RNA genomes, *Science* **215**:1577-1585.

Horiuchi, K., 1975, Genetic studies of RNA phages, in: *"RNA Phages"*, N. D. Zinder, ed., Cold Spring Harbor Laboratory, New York, pp. 29-50.

Jennings, P.A., Finch, J.T., Winter, G., and Robertson, J.S., 1983, Does the higher order structure of the influenza virus ribonucleoprotein guide sequence rearrangements in influenza viral RNA?, *Cell* **34**:619-627.

Joshi, R. L., Joshi, S., Chapeville, F., and Haenni, A.L., 1983a, Primary and secondary structures of the tRNA-like regions of the genomes of plant RNA viruses, in: *"Endocytobiology II"*, H. E. A. Schenk, and W. Schwemmler, eds., Walter de Gruyter, Berlin, pp. 57-68.

Joshi, S., Joshi, R.L., Haenni, A.L., and Chapeville, F., 1983b, tRNA-like structures in genomic RNAs of plant viruses, *Trends Biochem. Sci.* **8**:402-404.

Kamer, G., and Argos, P., 1984, Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses, *Nucl. Acids Res.* **12**:7269-7282.

Keck, J. G., Matsushima, G .K., Makino, S., Fleming, J .O., Vannier, D. M., Stohlman, S. A., and Lai, M. M. C., 1988, In vivo RNA-RNA recombination of coronavirus in mouse brain, *J. Virol.* **62**:1810-1813.

King, A. M. Q., McCahon, D., Slade, W. R., and Newman, J. W. I., 1982, Recombination in RNA, *Cell* **29**:921-928.

King, A. M. Q., Ortlepp, S. A., Newman, J. W. I., and McCahon, D., 1987, Genetic recombination in RNA viruses, in: *"The Molecular Biology of the Positive Strand RNA Viruses"*, D. J. Rowlands, M. A. Mayo, and B. W. J. Mahy, eds., Academic Press, London, pp. 129-152..

Kirkegaard, K., and Baltimore, D., 1986, The mechanism of RNA recombination in poliovirus, *Cell* **47**:433-443.

Kolakofsky,D., and Roux,L., 1987, The molecular biology of paramyxoviruses, in: *"The Molecular Basis of Viral Replication"*, R. Perez Bercoff, ed., Plenum Press, New York, London, pp. 277-297.

Konarska, M. M., Padgett, R. A., and Sharp, P. A., 1985, *Trans* splicing of mRNA precursors in vitro, *Cell* **42**:165-171.

Kuge, S., Saito, I, and Nomoto, A., 1986, Primary structure of poliovirus defective-interfering particle genomes and possible generation mechanisms of the particles, *J.Mol.Biol.* **192**:473-487.

Lai, M. M. C., Baric, R. S., Makino, S., Keck, J. G., Egbert, J., Leibowitz, J. L., and Stohlman, S. A., 1985, Recombination between nonsegmented RNA genomes of murine coronaviruses, *J. Virol.* **56**:449-456.

Lane, D., 1988, Enlarged family of putative helicases, *Nature* **334**:478.

Lazzarini, R. A., Keene, J. D., and Schubert, M., 1981, The origins of defective interfering particles of the negative-strand RNA viruses,*Cell* **26**:145-154.

Ledinko, N., 1963, Genetic recombination with poliovirus type 1. Studies of crosses between a normal horse serum-resistant mutant and several guanidine-resistant mutants of the same strain, *Virology* **20**:107-119.

Makino, S., Keck, J. G., Stohlman, S. A., and Lai, M. M. C., 1986a, High-frequency RNA recombination of murine coronaviruses, *J. Virol.* **57**:729-737.

Makino, S., Stohlman, S. A., and Lai, M. M. C., 1986b, Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription, *Proc. Natl. Acad. Sci. USA* **83**:4204-4208.

Mayo, M. A., 1987, A comparison of the translation strategies used by bipartite genome, RNA plant viruses, in: *"The Molecular Biology of the Positive Strand RNA Viruses"*, D. J. Rowlands, M. A. Mayo, and B. W. J. Mahy, eds., Academic Press, London, pp. 177-205.

McCahon, D., and Slade, W. R., 1981, A sensitive method for the detection and isolation of recombinants of foot-and-mouth disease virus, *J. gen. Virol.* **53**:333-342.

Miller, W. A., Waterhouse, P. M., and Gerlach, W. L., 1988, Sequence and organization of barley yellow dwarf virus genomic RNA, *Nucl. Acids Res.* **16**:6097-6111.

Monroe, S. S., and 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**:3279-3283.

Monroe, S. S., and Schlesinger, S., 1984, Common and distinct regions of defective-interfering RNAs of Sindbis virus, *J. Virol.* **49**:825-872.

Morch, M. D., and Haenni, A.L., 1987, Organization of plant virus genomes that comprise a single RNA molecule, in: *"The Molecular Biology of the Positive Strand RNA Viruses"*, D. J. Rowlands, M. A. Mayo, and B. W. J. Mahy, eds., Academic Press, London, pp. 153-175.

Morch, M. D., Boyer, J. C., and Haenni, A. L., 1988, Overlapping open reading frames revealed by complete nucleotide sequencing of turnip yellow mosaic virus genomic RNA, *Nucl. Acids Res.* **16**:6157-6173.

Morch, M. D., Valle, R.P.C., and Haenni, A.L., 1987, Regulation of translation of viral mRNAs, in: *"The Molecular Basis of Viral Replication"*, R. Perez Bercoff, ed., Plenum Press, New York, London, pp. 113-159.

Murant, A. F., and Mayo, M. A., 1982, Satellites of plant viruses, *Ann. Rev. Phytopathol.* **20**:49-70.

Nishihara, T., Mills, D. R., and Kramer, F. R., 1983, Localization of the Qβ replicase recognition site in MDV-1 RNA, *J. Biochem.* **93**:669-674.

Perrault, J., 1981, Origin and replication of defective interfering particles, *Curr. Top. Microbiol. Immunol.* **93**:151-207.

Plotch, S. J., Bouloy, M., Ulmanen, I., and Krug, R. M., 1981, A unique cap (m^7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription, *Cell* **23**:847-858.

Pringle, C. R., 1965, Evidence of genetic recombination in foot-and-mouth disease virus, *Virology* **25**:48-54.

Reanney, D. C., 1982, The evolution of RNA viruses, *Ann. Rev. Microbiol.* **36**:47-73.

Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B., and Vriend, G., 1985, Structure of a human common cold virus and functional relationship to other picornaviruses, *Nature* **317**:145-153.

Simon, A. E., and Howell, S. H., 1986, The virulent satellite RNA of turnip crinkle virus has a major domain homologous to the 3' end of the helper virus genome, *EMBO J.* **5**:3423-3428.

Söderlund, H., Keränen, S., Lehtovaara, P., Palva, I., Pettersson, R. F., and Kääriäinen, L.,1981, Structural complexity of defective-interfering RNAs of Semliki Forest virus as revealed by analysis of complementary DNA.*Nucl. Acids Res.* **9**:3403-3417.

Solnick, D., 1985, *Trans* splicing of mRNA precursors, *Cell* **42**:157-164.

Strauss, J. H., and Strauss, E. G., 1988, Evolution of RNA viruses, *Ann. Rev. Microbiol.* **42**: in press.

Strauss, J. H., Strauss, E. G., Hahn, C. S., Hahn, Y. S., Galler, R., Hardy, W. R., and Rice, C. M., 1987, Replication of alphaviruses and flaviviruses: proteolytic processing of polyproteins, in: *"Positive Strand RNA Viruses"*, M. A. Brinton, and R. R. Rueckert, eds., Alan R. Liss, Inc., New York, pp. 209-225.

Stollar, V., 1980, Defective interfering alphaviruses, in: *"The Togaviruses, Biology, Structure, Replication"*, R. W. Schlesinger, ed., Academic Press, New York, pp. 427-457.

Symons, R. H., Haseloff, J., Visvader, J. E., Keese, P., Murphy, P. J., Gordon, K. H. J., and Bruening, G., 1985, On the mechanism of replication of viroids, virusoids and satellite RNAs, in: *"Subviral Pathogens of Plants and Animals: Viroids and Prions"*, K. Maramorosch, and J. J. McKelvey, eds., Academic Press, New York, pp. 235-263.

Takamatsu, N., Ohno, T., Meshi, T., and Okada, Y., 1983, Molecular cloning and nucleotide sequence of the 30K and the coat protein cistron of TMV (tomato strain) genome, *Nucl. Acids Res.* **11**:3767-3778.

Taylor, M., W., and Hershey, H., V., 1987, Viruses: an overview, in: *"The Molecular Basis of Viral Replication"*, R. Perez Bercoff, ed., Plenum Press, New York, London, pp. 3-23.

Tsiang, M., Monroe, S.S., and Schlesinger, S., 1985, Studies of defective interfering RNAs of Sindbis virus with and without $tRNA^{Asp}$ sequences at their 5' termini, *J.Virol.* **54**:38-44.

Tsiang, M., Weiss, B. G., and Schlesinger, S., 1988, Effects of 5'-terminal modifications on the biological activity of defective interfering RNAs of Sindbis virus, *J.Virol.* **62**:47-53.

Van Duin, J., 1988, Single-stranded RNA bacteriophages, in: *"The Bacteriophages"*, R. Calendar, ed., Plenum Press, New York, vol. **1**, pp. 117-167.

Van Vloten-Doting, L., and Jaspars, E. M. J., 1977, Plant covirus systems: three-component systems, in: *"Comprehensive Virology"*, H. Fraenkel-Conrat, and R. R. Wagner, eds., Plenum Press, New York, vol. **11**, pp. 1-53.

Webster, R. G., Laver, W. G., Air, G. M., and Schild, G. C., 1982, Molecular mechanisms of variation in influenza viruses, *Nature* **296**:115-121.

Wellink, J., and Van Kammen, A., 1988, Proteases involved in the processing of viral polyproteins, *Arch. Virol.* **98**:1-26.

Zimmern, D., 1982, Do viroids and RNA viruses derive from a system that exchanges genetic information between eukaryotic cells?, *Trends Biochem. Sci.* **7**:205-207.