VIROLOGY 235, 1–9 (1997)
ARTICLE NO. VY978681

MINIREVIEW

New Insights into the Mechanisms of RNA Recombination

Peter D. Nagy and Anne E. Simon[1]

Department of Biochemistry and Molecular Biology, University of Massachusetts, Amherst, Massachusetts 01003

Received May 15, 1997; returned to author for revisions June 3, 1997; accepted June 16, 1997

RNA plays many different roles in biological systems. One of RNAs most intriguing features is the ability to carry genetic information despite a labile nature. While RNA genomes were presumably widespread in the primordial "RNA world," only RNA viruses and viroids continue to exploit RNA-based genetic materials. As with DNA-based organisms, these entities must evolve and adapt in order to survive. Genetic recombination, the formation of chimeric molecules from segments previously separated on the same molecule or present on different parental molecules, is one of the most important mechanisms for generating novel genomes that may have selective advantages over parental genomes. Recent studies on the evolution of RNA viruses have revealed that RNA recombination is a widespread phenomenon that has shaped modern viruses by rearranging viral genomes or disseminating functional modules among different viruses (Strauss and Strauss, 1988; Dolja and Carrington, 1992; Lai, 1992; Simon and Bujarski, 1994). In addition, an important short-term function of genetic recombination may be the rescue of functional sequences from mutated parental molecules, which is of particular significance given the high mutation rates associated with replication by RNA-dependent RNA polymerases (RdRp) (Domingo *et al.,* 1996). Depending on the precision of the repair mechanism, the repaired genome can be similar to the parental genome, or it can contain further mutations. This illustrates that sequence diversity in RNA sequences generated by genetic recombination can involve both gross changes and minor mutations. Moreover, the products of recombination can overlap with unrelated phenomena, including mutagenesis and nonrecombination-based genome repair.

Early studies identified RNA recombination events indirectly by isolating chimeric products that contained characteristic marker mutations from the parental RNAs. Following the initial recombination event, detection of recombinants required amplification in cells that imposed a selection for or against a particular recombinant based on its ability to be amplified *in vivo.* Nevertheless, these data were used to group RNA recombination as homologous or nonhomologous (King, 1988). Homologous recombinants were those derived from parental RNAs that were either very similar in the case of intratypic recombination or somewhat less similar in the case of intertypic recombination. Based on the precision of the event, Lai (1992) further divided this grouping of recombinants into homologous or aberrant homologous recombinants. Homologous recombinants were newly defined as containing no sequence alterations (except the existing marker mutations) when compared to the parental molecules while aberrant homologous recombinants contained modifications such as mismatch mutations, deletions, or insertions at or close to the junction site. Recombinants were grouped as nonhomologous if they were generated by recombination between dissimilar viral genomes or between a viral genome and host RNA. This classical genetic grouping of recombinants, however, did not take into account the mechanisms leading to their generation. Not surprisingly, recent studies have suggested similar recombinants can be formed by different mechanisms. Moreover, similar mechanisms have been postulated that lead to the formation of some recombinants that currently would be grouped as homologous or nonhomologous.

The most accepted models of RNA recombination are the replicase-driven template switching model, the RNA breakage and ligation model, and the breakage-induced template switching model. Since all these recombination models can lead to the formation of similar recombinant RNAs, it is not possible to promote a particular mechanism based only on the sequence of the recombination end-products. Rather, characterization of recombination intermediates and components of the recombination machinery is required. Below, we will summarize our current

[1] To whom correspondence and reprint requests should be addressed to at Department of Biochemistry and Molecular Biology, Lederle Graduate Research Center, University of Massachusetts, Box 34505, Amherst, MA 01003-4505. Fax: 413-545-4529; E-mail: simon@biochem.umass.edu.

0042-6822/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

knowledge on the molecular aspects of RNA recombination. Based on these data, we will propose a new classification system for recombinants based on both the mechanism of their generation and the nature of the recombinant product.

REPLICASE-DRIVEN TEMPLATE SWITCHING MECHANISM

In the replicase-driven template switching model, the RNA templates and viral replicase (RNA-dependent RNA polymerase, RdRp) are vital components of RNA recombination. At least three RNAs are involved in recombination events: the primary RNA template (donor strand), the nascent strand synthesized from the primary RNA template, and the acceptor strand.

The function of the donor strand (in conjunction with the nascent strand) in the template switching reaction is to halt synthesis of the nascent strand temporarily, thus providing an opportunity for the RdRp and/or the nascent strand to interact with the acceptor RNA, leading to template switching events. Intrinsic signals present on the donor or nascent RNAs are of two types: (1) pausing (or arrest) signals, which can halt the RdRp, but from which it can escape spontaneously, and (2) terminators, which can release the RdRp from the RNAs. Pausing or termination of viral RdRps, mediated by the sequence and/or secondary structure of the donor or nascent RNA, may occur similar to that of DNA-dependent RNA polymerases and RNA-dependent DNA polymerases (reverse transcriptases, RT; Wu *et al.,* 1995). For example, short U-rich stretches 3′ of stable hairpin structures in the nascent strand promote transcription termination by T7 bacteriophage and *Escherichia coli* RNA polymerases (Macdonald *et al.,* 1993; Wilson and von Hippel, 1995). In addition, specific terminator proteins bound to terminator sequences immediately downstream of U-rich regions are integral parts of the RNA polymerase I system in eukaryotes (Lang *et al.,* 1994; Lang and Reeder, 1995). The 5′ end of RNA templates promote RT-mediated template switching, possibly by causing pausing and/or termination (Peliska and Benkovic, 1992, 1994). Template switching involving the 5′ end of DNA templates were also observed for *E. coli* RNA polymerase *in vitro* (Nudler *et al.,* 1996). Strong hairpin structures present on either the template or nascent RNA were pausing signals for Qβ RdRp (Mills *et al.,* 1978) and detection of RNA transcripts of discrete sizes was likely due to RdRp pausing at strong hairpin structures in coronavirus infections (Baric *et al.,* 1987).

The data described above suggest that pausing and termination are characteristic features of all RNA polymerases, which can, in turn, promote template switching. Thus one prediction of a template switching model is that regions in RNA that promote RdRp pausing or termination will constitute recombination hotspots. Accordingly, the end of sat-RNA D [a naturally occurring subviral RNA associated with turnip crinkle virus (TCV)] constitutes one of the recombination hotspots in the TCV-based recombination system (Cascone *et al.,* 1990, 1993; Carpenter *et al.,* 1995). Similarly, artificially created 5′ ends and stable internal hairpin structures that may promote replicase pausing are found at the recombination junction sites in a tombusvirus system (White and Morris, 1995). A role for replicase pausing in template switching was also proposed for recombination in brome mosaic virus (BMV). Regions capable of forming stable heteroduplexes were introduced into BMV RNAs and recombination was targeted to these sequences; such an extended heteroduplex region may temporarily impede the progression of the RdRp before unwinding the heteroduplex leads to the resumption of rapid RNA elongation (Fig. 1A; Nagy and Bujarski, 1993; Nagy *et al.,* 1995). Indeed, weakening the stability of the heteroduplex by introducing mismatched regions into the heteroduplex caused a shift in recombination junction sites toward more stable portions of the heteroduplex. Moreover, the association of recombination with heteroduplexes suggests that the donor RNA participates in template switching by forcing the RdRp to pause and by helping to bring the acceptor RNA into proximity (Fig. 1A). A role for similar heteroduplexes in promoting replicase pausing and template switching but formed from intramolecular base-pairing has been implicated in the formation of tombusvirus (White and Morris, 1995) and bromovirus DI RNAs (Romero *et al.,* 1993; Pogany *et al.,* 1995; Pogany, 1997).

A/U-rich sequences may also promote recombination by inducing replicase pausing in a homologous recombination system in BMV (Fig. 1C; Nagy and Bujarski, 1995, 1996) and in an intramolecular recombination system in picornaviruses (Pilipenko *et al.,* 1995). Such A/U-rich or U-rich sequences are thought to promote RdRp slippage resulting in the accidental incorporation of nontemplated nucleotides at the 3′ end of the growing nascent strand. This in turn may force replicase pausing and promote template switching. Nontemplated nucleotides were frequently observed at the junction sites in TCV (Cascone *et al.,* 1990, 1993; Carpenter *et al.,* 1995), tobamovirus (Raffo and Dawson, 1991) and BMV recombinants (Nagy and Bujarski, 1993, 1996). A somewhat more speculative role for nontemplate sequences in replicase pausing and template switching has also been proposed for alphaviruses (Raju *et al.,* 1995).

The role of the acceptor RNA during template switching events is likely to be complex. The RdRp must be able to bind the acceptor RNA and use the 3′ end of the nascent RNA as a primer during the reinitiation of RNA synthesis. The role of the acceptor RNA in template switching events is best described for recombination between TCV satellite RNAs. A stem–loop structure (designated motif1-hairpin) on the sat-RNA C acceptor RNA is required for recombination to occur at the 3′ base of the

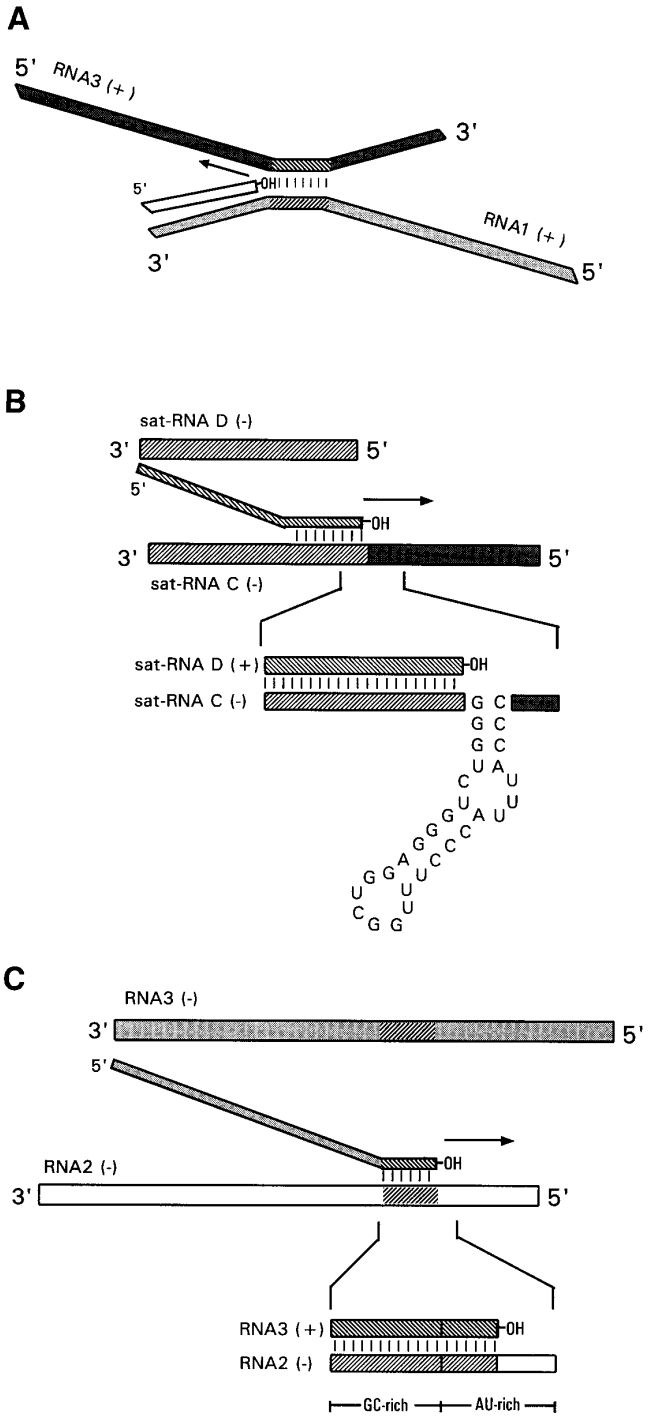

FIG. 1. Current replicase-mediated template-switching models of RNA recombination in TCV and BMV. (A) Heteroduplex-mediated recombination between positive strands of BMV RNA1 and RNA3. (B) Recombination between satellite RNAs associated with TCV. The sequence of the required motif1-hairpin is shown. (C) Recombination events within the identical regions of BMV RNA2 and RNA3. Recombination is favored when GC-rich and AU-rich sequences are located as shown.

hairpin *in vivo* (Fig. 1B; Cascone *et al.,* 1993). Recent data obtained with an *in vitro* (cell-free) system also demonstrated an important role for the motif1-hairpin in a reaction that mimics *in vivo* recombination (P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation). The template for the *in vitro* system is a chimeric RNA containing the recombination hot-spot from sat-RNA D plus-strands at the 3′ end connected by a short loop to sat-RNA C minus-strands including the motif-1 hairpin (see Fig. 1B for the *in vivo* equivalent). The motif1-hairpin was found to be essential for efficient synthesis of panhandle-like molecules generated by intramolecular extensions from the 3′ end of the chimeric template in the vicinity of the hairpin (analogous to a primer extension reaction) (P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation). Importantly, a role for a short intrastrand base-paired region between the sat-RNA D sequence at the 3′ end of the chimeric RNA and a region just upstream of the motif1-hairpin was found in the *in vitro* system. Testing of corresponding constructs *in vivo* for recombination at the motif1-hairpin also suggested an important role for base-pairing between the nascent strand and the acceptor strand 3′ of the junction site (Fig. 1B; P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation). Based on these studies and others indicating that the motif1-hairpin shares sequence similarity with mapped TCV promoters (Cascone *et al.,* 1990; H. Guan, C. Song, and A. E. Simon, manuscript in preparation), two separate domains on the acceptor TCV satellite RNA have roles in template switching. One domain 3′ of the hairpin (on sat-RNA C minus-strands) serves to bind the nascent RNA, while the second domain, the motif-1 hairpin, is postulated to bind the replicase (Fig. 1B; P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation). Short sequence complementarity between the nascent RNA and the acceptor RNA was also postulated to facilitate the association of nascent strands with acceptor strands in BMV (Nagy and Bujarski, 1995), nodavirus (Li and Ball, 1993), tombusvirus (White and Morris, 1995), tobamovirus (Raffo and Dawson, 1991), Qβ, MS2, and ϕ6 bacteriophage (Biebricher and Luce, 1992; Onodera *et al.,* 1993; Olsthoorn and van Duin, 1996; Qiao *et al.,* 1997) recombinants, and tombusvirus (White and Morris, 1994a; 1994b), mycovirus (Shapira *et al.,* 1991), and bromovirus (Pogany *et al.,* 1995; Pogany, 1997) DI RNAs. Homologous recombination was also observed between long stretches of homologous sequences in poliovirus (Kirkegaard and Baltimore, 1986; Romanova *et al.,* 1986; King, 1988; Jarvis and Kirkegaard, 1992), coronavirus (Makino *et al.,* 1987; Koetzner *et al.,* 1992; van der Most *et al.,* 1992), alphavirus (Raju *et al.,* 1995), and Qβ bacteriophage (Palasingam and Shaklee, 1992). Annealing between the nascent and the acceptor RNA strands can be frequently imprecise in A/U-rich regions, leading to sequence deletions or insertions in BMV and picornavirus recombinants (Nagy and Bujarski, 1995, 1996; Pilipenko *et al.,* 1995). Similar imprecise recombinants were frequently detected in an *in vitro* RT-based recombination system (Wu *et al.,* 1995).

The most notable role of the nascent RNA strand during template switching is its function as a primer. Extension of the 3′ end of the nascent strand by the replicase using the acceptor RNA as template results in the generation of the recombinant RNA. In contrast to DNA polymerases and RTs that favor primers that contain a 3′-terminal nucleotide able to base-pair with the template, some viral RdRps appear to be able to extend primers with or without a base-paired 3′ end. This was experimentally demonstrated for TCV in an *in vitro* 3′ end extension system in which the TCV RdRp readily extended from the 3′ terminus of primers located proximal to the motif1-hairpin even though the ultimate and penultimate bases were not paired with the template (Fig. 1B; P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation).

In addition to its primer function, the nascent RNA likely plays a role in RdRp pausing and in selection of the reinitiation (junction) site. The nascent RNA in coronavirus and poliovirus recombination is also thought to be processed by a cleavage reaction before association with the acceptor RNA (Lai, 1992). Processing of the 3′ end of the nascent strand before template switching may also explain why many of the sites were asymmetric in heteroduplex-mediated recombination (i.e., the donor and acceptor sites were located far apart on the heteroduplex) in BMV (Nagy and Bujarski, 1993; Nagy *et al.,* 1995). Processing the 3′ end of the nascent strand within a pausing polymerase complex has been described for *E. coli* RNA polymerase, eukaryotic RNA polymerase II, and vaccinia virus RNA polymerase and is facilitated by nonpolymerase proteins such as GreA, GreB, and TFIIS (Borukhov *et al.,* 1993; Hagler and Schuman, 1993; Wang and Hawley, 1993).

In addition to the above described three core RNA components of recombination systems, other RNA molecules can potentially influence recombination frequency and junction site selection. For example, negative-stranded viral RNA may facilitate bringing the donor and acceptor RNAs of poliovirus into proximity (Kuge *et al.,* 1986). Positioning and strengthening of less stable heteroduplexes formed between the parental RNAs may be achieved by their negative-stranded counterparts in BMV (Bujarski *et al.,* 1994). Overall, the above examples demonstrate that the RNA components of the recombination complex can influence many characteristics of recombination, including frequency, junction site selection, and precision.

How an RdRp switches template is not well known. We speculate that the properties of a given viral RdRp will influence each of the steps during the recombination event. For example, due to different processivities, RdRps may pause at different sequences and at variable frequencies. The more processive RdRps are predicted to recombine in a processive manner (Jarvis and Kirkegaard, 1991), while less processive RdRps may favor a nonprocessive pathway. For example, a processive model was proposed for heteroduplex-mediated recombination in BMV (Nagy and Bujarski, 1993; Bujarski and Nagy, 1994, 1996). This model suggested that the RdRp approaching the heteroduplex region pauses and then switches to the nearby sequence of the acceptor RNA along with the nascent strand (Fig. 1A). A similar "looping out" mechanism may lead to the formation of some DI RNAs in tombusvirus and bromoviruses (Pogany *et al.,* 1995; White and Morris, 1995; Pogany, 1997). A different type of processive model proposed for poliovirus and bromovirus homologous recombination predicts that the pausing RdRp slides backwards 10 to 20 nt on the acceptor template (Jarvis and Kirkegaard, 1991; Nagy and Bujarski, 1995, 1996) resulting in the extrusion of the 3′ terminus of the nascent strand from the RdRp while sequences just 5′ are still bound by the RdRp. The exposed 3′ terminus of the nascent strand would then hybridize to a complementary region on the acceptor template. Finally, the RdRp slides forward and reinitiates nascent strand synthesis on the acceptor RNA.

The nonprocessive model, proposed for TCV (Cascone *et al.,* 1990, 1993) and coronavirus (Lai, 1992) recombinants, suggests that the pausing RdRp dissociates from the primary template, remaining either in association with the nascent strand or releasing the nascent strand, followed by reassociation of the RdRp and nascent RNA with the acceptor RNA. This nonprocessive model is applicable to most of the replicase-driven recombination systems and the RT-driven recombination systems (Hu and Temin, 1990). One of the most important questions for nonprocessive model is, how does a promoter-dependent RdRp recognize and associate with nonpromoter sequences on the acceptor RNA prior to resumption of nascent strand synthesis? Although poliovirus RdRp and T7 bacteriophage and *E. coli* RNA polymerases are readily able to extend on randomly hybridizing primers *in vitro* (Daube and von Hippel, 1992; Paul *et al.,* 1994; Triana-Alonso *et al.,* 1995), partially purified BMV and TCV RdRps cannot extend on primers that anneal at random internal positions (Kao and Sun, 1996; Nagy *et al.,* 1997; C. Kao, personal communication; P. D. Nagy and A. E. Simon, unpublished). However, BMV and TCV RdRps are able to extend on short primers complementary to 3′-terminal promoter sequences on their plus-strand templates, suggesting that extension of primers by these RdRps requires a proximal promoter element (Kao and Sun, 1996; Nagy *et al.,* 1997). The nearby presence of a hairpin containing features of TCV promoters was required for recombination *in vivo* between satellite RNAs of TCV, and recent *in vitro* results suggest that the TCV RdRp is recruited by the hairpin structure (P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation). Interestingly, similar to TCV, the unique promoter-dependent RT encoded by the Mauriceville plasmid can bind to specific internal positions of the template and initiate primer extension (Wang and Lambowitz, 1993). Subgenomic promoters for many viruses may also be used to recruit RdRps (Keck *et al.,* 1987; Allison *et al.,*

1990; Graves and Roossink, 1995; Miller *et al.,* 1995; C. D. Carpenter and A. E. Simon, unpublished). No such sequences or structures, however, exist for the BMV system at or close to the junction sites, suggesting that the BMV recombination machinery may favor the processive pathway described above.

Recent studies with BMV RdRp mutants revealed a role for the replicase component in the frequency and precision of template switching and in junction site selection. For example, mutations within the helicase-like domain of the BMV 1a replicase protein resulted in an increase in the frequency of recombination and shifted the recombination sites into energetically less stable portions of a heteroduplex that could form between the two parental RNAs (Nagy *et al.,* 1995). The progression of the replicase complex is thought to be impeded by mutations in the 1a protein, facilitating template switching at a higher frequency, and at different sites when compared with recombination using the wild-type replicase (Nagy *et al.,* 1995). In contrast to the 1a mutations, a mutations within the BMV 2a replicase protein decreased the percentage of recombination in BMV (Bujarski and Nagy, 1996; Figlerowitz *et al.,* 1997). Thus the frequency of recombination using the wild-type BMV RdRp can be both up and down regulated by various mutations. Moreover, mutations within the 1a and 2a proteins influenced the precision of recombination with the percentage of junctions containing nontemplated or mismatched nucleotides increasing for some of the mutants. Altogether, these experiments suggest that the different frequencies of recombination and DI RNA formation observed for various viruses might be the consequence of differential abilities of the viral replicases to mediate template switching.

RNA BREAKAGE AND LIGATION MECHANISM

Models for recombination (Lazzarini *et al.,* 1981; King, 1988) suggesting that chimeric RNAs are formed by a breakage and ligation mechanism are based on the well-characterized DNA-based breakage and ligation recombination systems. Although ribonuclease-mediated RNA scission and ligation mechanism was proposed previously (Tsagris *et al.,* 1991), it has not yet been formally demonstrated. In contrast, a site-specific ribozyme-mediated RNA breakage and ligation mechanism has been demonstrated for splicing of group II introns *in vitro* (Morl and Schmelzer, 1990) and is used to repair bacterial mRNA *in vivo* (Sullenger and Cech, 1994). Recently, a transesterification mechanism was proposed to explain the *in vitro* generation of recombinants between RNAs associated with Qβ bacteriophage (Chetverin *et al.,* 1997). The authors based their conclusions on the observation that altering the 3′-end OH group of the 5′ (acceptor) template interfered with recombinant isolation. In addition, recombination frequently occurred at the 3′ end of the 5′ template. Although these data can be explained by an RNA-mediated transesterification mechanism, they do not exclude a template switching mechanism since Qβ replicase was required to amplify the generated recombinants. One possibility is that the Qβ replicase requires the 3′ OH group of the 5′ template for binding and reinitiation of synthesis from the nascent strand. Furthermore, the bulky and charged side-groups introduced into the 3′ end of the acceptor RNA may have interfered with the binding of the Qβ replicase during template switching. Even if the recombinant products generated in the *in vitro* Qβ bacteriophage system are proven to be the result of an RNA-mediated breakage and ligation mechanism, recombination using the artificial sequences of this system has not yet been demonstrated *in vivo.* Therefore, it is difficult to estimate how widespread an RNA-mediated breakage and ligation mechanism might be in natural virus systems.

BREAKAGE-INDUCED TEMPLATE SWITCHING MECHANISM

The breakage-induced template switching model predicts that breakage of the donor RNA can promote replicase-driven template switching. This model is similar to the template switching model, with the additional requirement for cleavage of the donor RNA. The 5′ end of the RNA template at the cleavage site is predicted to cause replicase pausing (or termination) followed by template switching of the replicase. This is a plausible possibility due to the lability of RNA molecules and the strong pausing sites generated due to the presence of new 5′ termini. Indeed, artificially created 5′ ends in donor RNA that resemble cleaved or processed RNAs served as recombination hot-spots in tombusviruses (White and Morris, 1995). In addition, recombination mediated by RNA cleavage has been described for RT systems as well (Coffin, 1979; Peliska and Benkovic, 1992, 1994). RNA cleavage can be caused by exo- or endonucleolytic cleavages during natural degradation processes in cells or by ribozymes or metal ions. Due to the susceptibility of RNA to degradation, it is difficult to exclude that even seemingly "standard" template switching recombination events (see above) occur with cleaved donor RNA instead of full-length (input) RNA. Regardless of the mechanism of recombination, a requirement for *in vivo* or *in vitro* amplification can lead to selection for or against particular recombinants. Therefore, characterization of the role that selection plays in producing populations of recombinants in different systems should be kept in mind when junction sites are analyzed.

HOMOLOGOUS VERSUS NONHOMOLOGOUS RECOMBINATION: REVISITING CLASSICAL DEFINITIONS

The terms "homologous" and "nonhomologous" recombination introduced previously for grouping types of RNA re-

combination were adapted from DNA recombination. Due to basic differences between DNA and RNA recombination, these terms cannot be easily applied to RNA-based systems. The confusion in using such classical terms comes from the flexible nature of RNA recombination. For example, very short (5 to 15 nt) sequence similarity between parental RNAs of BMV facilitates RNA recombination. Such lengths of sequence similarity, however, can be found between otherwise dissimilar RNAs. The matter is further complicated by the finding that not all regions of sequence similarity (even 60 nt long) support recombination in BMV (Nagy and Bujarski, 1993, 1995). Short stretches of similarity (2 to 5 nt) have also been found in donor and acceptor RNAs at junction sites for some RNA recombinants and have been proposed to facilitate recombination (Raffo and Dawson, 1991; Shapira *et al.,* 1991; Biebricher and Luce, 1992; Li and Ball, 1993; Nagy and Bujarski, 1995; Pogany *et al.,* 1995; White and Morris, 1995; Qiao *et al.,* 1997). Are all these recombination events therefore of "homologous" type? The term "homologous" is also problematic, since it implies that the sequences in question have a common ancestry. Another confusion in the recombination literature is the different types of recombinants generated in the TCV system. For instance, are recombinant products obtained from template switching between two satellite RNAs that share sequence similarity in the region where recombination occurred "aberrant homologous" recombinants, while those between a satellite and the genomic RNAs of "nonhomologous" type (Lai, 1992)? Both of these TCV recombination systems require hairpin structures present only in the acceptor RNAs and have very similar donor junction sites (Cascone *et al.,* 1993; Carpenter *et al.,* 1995). These and other observations suggest that TCV recombinants are generated by a similar mechanism where hairpin structures are the major determinants of RNA recombination (Simon and Nagy, 1996). Why then is one type of TCV recombination classified as (aberrant) homologous while the second is nonhomologous?

To redefine recombination according to current theories, various criteria need to be examined including the structure of the intermediates, the recombination end products, and the recombination machinery. In line with previous definitions, recombinants will be grouped according to RNA elements on the parental (and not donor) RNAs. We propose that recombination be divided into three classes: Class 1 recombination is *similarity-essential* recombination, in which substantial sequence similarity between the parental RNAs is required and is the major RNA determinant of the recombination event (Fig. 2A). The end-products can be of two types: precise and imprecise (aberrant). Precise recombinants contain no sequence alterations within the region of sequence similarity when compared to the parental molecules. In contrast, imprecise recombinants contain diverged sequences within the region of sequence similarity, including mutations, deletions, insertions, and nontemplate nucleotides. Class 1 recombination events may occur frequently between identical parental RNAs and go unnoticed in the absence of marker mutations.

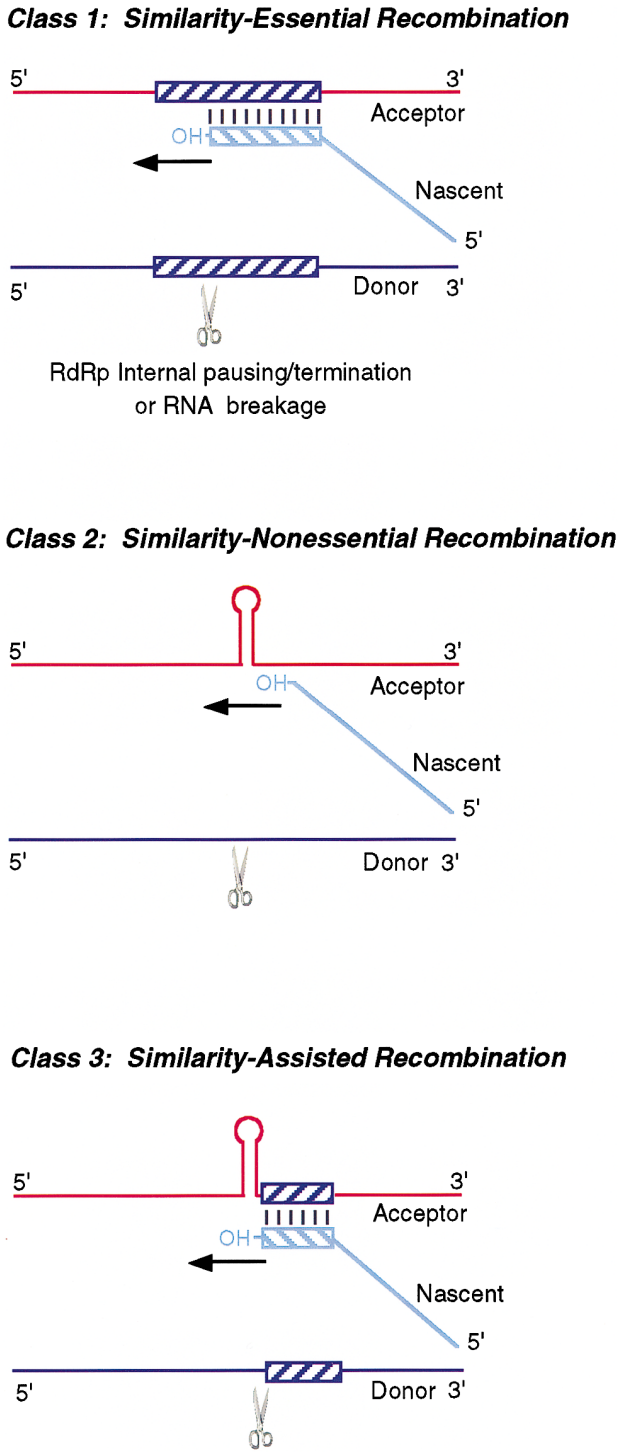

FIG. 2. Three classes of RNA recombination. Replicase-mediated RNA synthesis after the template-switch events is shown by an arrow. The hairpin structure shown on the acceptor RNA symbolically represents various RNA determinants that are required for Class 2 and Class 3 recombination.

According to the current models (see above), the role

of sequence similarity between the parental RNAs during the template-switching events is to facilitate annealing (base-pairing) between the nascent-strand and the acceptor RNA within the complementary region. Sequence similarty between the parental RNAs as a major RNA determinant of recombination is best exemplified in RT-based template-switching systems (Luo and Taylor, 1990; DeStefano *et al.,* 1992, 1994). The sequence requirements for this class of recombinants are best defined for BMV. Short (15–60 nt) sequence identity between two heterologous RNA components of BMV (RNA2 and RNA3) facilitated precise recombination events (Fig. 1C; Nagy and Bujarski, 1995). Decreasing the percentage similarity between RNAs 2 and 3 reduced the frequency of recombination, while mismatch mutations that increased the similarity between RNA3 and RNA1, caused a shift in recombination from RNA2 to RNA1 (Nagy and Bujarski, 1995). In addition, the sequence context (AU-rich versus GC-rich) of the similar sequences played a major role in determining the frequency and precision of the events (Nagy and Bujarski, 1996, 1997). Sequences flanking the hot-spot regions also influenced the frequency of recombination, demonstrating the complex nature of the recombination events (Nagy and Bujarski, submitted).

Class 2 recombination, or *similarity-nonessential* recombination, includes events where sequence similarity between the parental RNAs is not apparently required, although such similar regions may be present (Fig. 2B). Class 2 recombination would therefore include those events that do not require base-pairing between the nascent and acceptor strands. Other features of the RNAs, such as RdRp binding sequences, secondary structure elements, heteroduplex formation between parental RNAs, etc., determine the recombination characteristics. Sequence/structure requirements of Class 2 recombination events are best described for BMV, where sequence complementarity brings the parental RNAs into proximity and the formed heteroduplex between the RNAs influences the recombination frequency and junction site selection (Fig. 1A; Nagy and Bujarski, 1993; Nagy *et al.,* 1995). Although the driving force of recombination in this case is the formation of the heteroduplex, limited sequence similarity in the region of recombination may also affect the location of junction sites in some of the recombinants. The difference in mechanisms between Class 1 and Class 2 recombination can be seen for BMV, where a mutation in one subunit of the BMV replicase decreased the isolation of Class 2 recombinants, but not those of Class 1 (Figlerowitz *et al.,* 1997). Although Class 2 recombination probably includes many unrelated recombination pathways, currently there is insufficient data for placement of these pathways into separate groups. As more knowledge is gained concerning the mechanisms of recombination, these classifications will need to be revisited and modified.

Class 3 recombination or *similarity-assisted* recombination combines features of both Class 1 and Class 2 recombination. As with Class 1 recombination, sequence similarity between parental RNAs significantly influences the frequency or site of the recombination event. However, unlike Class 1 recombination, additional RNA determinants are essential for efficient recombination and such determinants are present only in one of the parental RNAs (Fig. 2C). An example of Class 3 recombination can be found in the above described TCV system, where regions of similar sequence between satellite RNAs influence the site and frequency of recombination events. In addition, the motif1-hairpin structure present only in the acceptor RNA is required for recombination and is involved in recruitment of the RdRp (P. D. Nagy, C. Zhang, and A. E. Simon, manuscript in preparation). While analysis of sequences at junction sites should allow for temporary placement of recombinants derived in various viral systems within these three classes, additional experimental analysis (e.g., mutagenesis of sequences at or near junction sites) is required for more definative classification.

In summary, there is no single mechanism that describes the variety of recombinants generated in different virus systems. Each component involved in recombination, whether RNA or protein, contributes to the observed differences. However, based on the assumption that many viral replicases, similar to RTs and *E. coli* RNA polymerase, are capable of template switching at some frequency, we believe that most RNA recombination events are mediated by the viral RdRps and RNA recombinants are byproducts of the replication machinery. It is possible that such byproducts interfere with the more important processes of viral replication since the frequency of RNA recombination is below its maximum possible level. This observation is based on the finding that small modifications in the RNAs or RdRp increase recombination frequencies in several viral systems (Nagy and Bujarski, 1993; 1997; Nagy *et al.,* 1995; White and Morris, 1995; Figlerowitz *et al.,* 1997; Pogany, 1997). In addition, RNA recombination that is too efficient may negatively impact on the integrity of the viral genomes, which, due to their quasispecies nature, already "live on the edge of catastrophe" (Domingo *et al.,* 1996; Nichol, 1996). More efficient RNA recombination could also lead to more frequent generation of parasitic RNAs, such as satellites, DIs, and chimeric RNAs, which frequently compete for replication with the viral genome (Roux *et al.,* 1991). In contrast, lack of recombination could decrease the ability of viruses to repair their genomes (Carpenter and Simon, 1996; Lai, 1996) or adapt to evolving hosts by not allowing viruses to incorporate evolution-tested functional modules or genes from other viruses or their hosts. Therefore, RNA recombination is like many things in nature: too much is as detrimental as too little.

ACKNOWLEDGMENTS

Research in the lab of A. E. Simon is supported by National Science Foundation awards MCB-9315948 and MCB-9630191. We are grateful

to L. A. Ball for critical reading of the manuscript and for very helpful suggestions.

REFERENCES

Allison, R. F., Thompson, G., and Ahlquist, P. (1990). Regeneration of a functional RNA virus genome by recombination between deletion mutants and requirement for cowpea chlorotic mottle virus 3a and coat genes for systemic infection. *Proc. Natl. Acad. Sci. USA* 87, 1820–1824.

Baric, R. S., Shieh, C.-K., Stohlman, S. A., and Lai, M. M. C. (1987). Analysis of intracellular small RNAs of mouse hepatitis virus: Evidence for discontinuous transcription. *Virology* 156, 342–354.

Biebricher, C. K., and Luce, R. (1992). *In vitro* recombination and terminal elongation of RNA by Qβ replicase. *EMBO J.* 11, 5129–5135.

Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from *E. coli*. *Cell* 72, 459–466.

Bujarski, J. J., and Nagy, P. D. (1994). Targeting of the site of nonhomologous recombination in brome mosaic virus. *Arch. Virol. (Suppl.)* 9, 231–238.

Bujarski, J. J., and Nagy, P. D. (1996). Different mechanisms of homologous and nonhomologous recombination in brome mosaic virus: Role of RNA sequences and replicase proteins. *Semin. Virol.* 7, 363–372.

Bujarski, J. J., Nagy, P. D., and Flasinski, S. (1994). Molecular studies of genetic RNA–RNA recombination in brome mosaic virus. *Adv. Virus Res.* 43, 275–302.

Carpenter, C. D., and Simon, A. E. (1996). *In vivo* repair of deletions at the 3′ end of a TCV satellite RNA *in vivo* may involve two abortive synthesis and priming events. *Virology* 226, 153–160.

Carpenter, C. D., Oh, J.-W., Zhang, C., and Simon, A. E. (1995). Involvement of a stem–loop structure in the location of junction sites in viral RNA recombination. *J. Mol. Biol.* 245, 608–622.

Cascone, P. J., Carpenter, C. D., Li, X.-H., and Simon, A. E. (1990). Recombination between satellite RNAs of turnip crinkle virus. *EMBO J.* 9, 1709–1715.

Cascone, P. J., Haydar, T. F., and Simon, A. E. (1993). Sequences and structures required for recombination between virus-associated RNAs. *Science* 260, 801–805.

Chetverin, A. B., Chetverina, H. V., Demidenko, A. A., and Ugarov, V. I. (1997). Nonhomologous RNA recombination in a cell-free system: Evidence for a transesterification mechanism guided by the secondary structure. *Cell* 88, 503–513.

Coffin, J. M. (1979). Structure, replication, and recombination of retrovirus genomes: Some unifying hypotheses. *J. Gen. Virol.* 42, 1–26.

Daube, S. S., and Von Hippel, P. H. (1992). Functional transcription elongation complexes from synthetic RNA–DNA bubble duplexes. *Science* 258, 1320–1324.

DeStefano, J. J., Mallaber, L. M., Rodriguez-Rodriguez, L., Fay, P. J., and Bambara, R. A. (1992). Requirements for strand transfer between internal regions of heteropolymeric RNA templates by human immunodeficiency virus reverse transcriptase. *J. Virol.* 66, 6370–6378.

DeStefano, J. J., Bambara, R. A., and Fay, P. J. (1994). The mechanism of human immunodeficiency virus reverse transcriptase catalyzed strand transfer from internal regions of heteropolymeric RNA templates. *J. Biol. Chem.* 269, 161–168.

Dolja, V. V., and Carrington, J. C. (1992). Evolution of positive-strand RNA viruses. *Semin. Virol.* 3, 315–326.

Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S. F., Quer, J., Novella, I. S., and Holland, J. J. (1996). Basic concepts in RNA virus evolution. *FASEB J.* 10, 859–864.

Figlerowitz, M., Nagy, P. D., and Bujarski, J. J. (1997). A mutation in the putative RNA polymerase gene inhibits nonhomologous, but not homologous, genetic recombination in an RNA virus. *Proc. Natl. Acad. Sci. USA* 94, 2073–2078.

Graves, M. V., and Roossink, M. J. (1995). Characterization of defective RNAs derived from the Fny strain of cucumber mosaic virus. *J. Virol.* 69, 4746–4751.

Hagler, J., and Shuman, S. (1993). Nascent RNA cleavage by purified ternary complexes of vaccinia RNA polymerase. *J. Biol. Chem.* 268, 2166–2173.

Hu, W.-S., and Temin, H. M. (1990). Retroviral recombination and reverse transcription. *Science* 250, 1227–1233.

Jarvis, T. C., and Kirkegaard, K. (1991). The polymerase in its labyrinth: Mechanisms and implications of RNA recombination. *Trends Genet.* 7, 186–191.

Jarvis, T. C., and Kirkegaard, K. (1992). Poliovirus RNA recombination: mechanistic studies in the absence of selection. *EMBO J.* 11, 3135–3145.

Kao, C. C., and Sun, J.-H. (1996). Initiation of minus-strand RNA synthesis by the brome mosaic virus RNA-dependent RNA polymerase: Use of oligoribonucleotide primers. *J. Virol.* 70, 6826–6830.

Keck, J. G., Stohlman, S. A., Soe, L. H., Makino, S., and Lai, M. M. C. (1987). Multiple recombination sites at the 5′-end of murine coronavirus RNA. *Virology* 156, 331–341.

King, A. M. Q. (1988). Genetic recombination in positive strand RNA viruses. *In* "RNA Genetics" (E. Domingo, J. J. Holland, and P. Ahlquist, Eds.), Vol. II, pp. 149–185. CRC Press, Boca Raton, FL.

Kirkegaard, K., and Baltimore, D. (1986). The mechanism of RNA recombination in Poliovirus. *Cell* 47, 433–443.

Koetzner, C. A., Parker, M. M., Ricard, C. S., Sturman, L. S., and Masters, P. S. (1992). Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. *J. Virol.* 66, 1841–1848.

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. (1992). RNA recombination in animal and plant viruses. *Microbiol. Rev.* 56, 61–79.

Lai, M. M. C. (1996). Recombination in large RNA viruses: coronaviruses. *Semin. Virol.* 7, 381–388.

Lang, W. H., Morrow, B. E., Ju, Q., Warner, J. R., and Reeder, R. H. (1994). A model for transcription termination by RNA polymerase I. *Cell* 79, 527–534.

Lang, W. H., and Reeder, R. H. (1995). Transcription termination of RNA polymerase I due to a T-rich element interacting with Reb1p. *Proc. Natl. Acad. Sci. USA* 92, 9781–9785.

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.

Li, Y., and Ball, L. A. (1993). Nonhomologous RNA recombination during negative-strand synthesis of flock house virus RNA. *J. Virol.* 67, 3854–3860.

Luo, G., and Taylor, J. M. (1990). Template switching by reverse transcriptase during DNA synthesis. *J. Virol.* 64, 4321–4328.

Macdonald, L. E., Zhou, Y., and McAllister, W. T. (1993). Termination and slippage by bacteriophage T7 RNA polymerase. *J. Mol. Biol.* 232, 1030–1047.

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

Miller, W. A., Dinesh-Kumar, S. P., and Paul, C. P. (1995). Luteovirus gene expression. *Crit. Rev. Plant Sci.* 14, 179–211.

Mills, D. R., Dobkin, C., and Kramer, F. R. (1978). Template-determined, variable rate of RNA chain elongation. *Cell* 15, 541–550.

Morl, M., and Schmelzer, C. (1990). Group II intron RNA-catalyzed recombination of RNA *in vitro*. *Nucleic Acids Res.* 18, 6545–6551.

Nagy, P. D., and Bujarski, J. J. (1993). Targeting the site of RNA–RNA recombination in brome mosaic virus with antisense sequences. *Proc. Natl. Acad. Sci. USA* 90, 6390–6394.

Nagy, P. D., and Bujarski, J. J. (1995). Efficient system of homologous RNA recombination in brome mosaic virus: Sequence and structure requirements and accuracy of crossovers. *J. Virol.* 69, 131–140.

Nagy, P. D., Dzianott, A., Ahlquist, P. G., and Bujarski, J. J. (1995). Mutations in the helicase-like domain of protein 1a alter the sites of RNA–RNA recombination in brome mosaic virus. *J. Virol.* 69, 2547–2556.

Nagy, P. D., and Bujarski, J. J. (1996). Homologous RNA recombination in brome mosaic virus: AU-rich sequences decrease the accuracy of crossovers. *J. Virol.* 70, 415–426.

Nagy, P. D., and Bujarski, J. J. (1997). Engineering of homologous recombination hotspots with AU-sequences in brome mosaic virus. *J. Virol.* 71, 3799–3810.

Nagy, P. D., Carpenter, C. D., and Simon, A. E. (1997). A novel 3′-end repair mechanism in an RNA virus. *Proc. Natl. Acad. Sci. USA* 94, 1113–1118.

Nichol, S. (1996). Life on the edge of catastrophe. *Nature* 384, 218–219.

Nudler, E., Avetissova, E., Markovtsov, V., and Goldfarb, A. (1996). Transcription processivity: Protein–DNA interactions holding together the elongation complex. *Science* 273, 211–217.

Olsthoorn, R. C. L., and van Duin, J. (1996). Random removal of inserts from an RNA genome: Selection against single-stranded RNA. *J. Virol.* 70, 729–736.

Onodera, S., Qiao, X., Gottlieb, P., Strassman, J., Frilander, M., and Mindich, L. (1993). RNA structure and heterologous recombination in the double-stranded RNA bacteriophage ϕ6. *J. Virol.* 67, 4914–4922.

Palasingam, K., and Shaklee, P. N. (1992). Reversion of Qβ RNA phage mutants by homologous RNA recombination. *J. Virol.* 66, 2435–2442.

Paul, A. V., Cao, X., Harris, K. S., Lama, J., and Wimmer, E. (1994). Studies with poliovirus 3D-pol: Stimulation of poly(U) synthesis *in vitro* by purified poliovirus protein 3AB. *J. Biol. Chem.* 269, 29173–29181.

Peliska, J. A., and Benkovic, S. J. (1992). Mechanism of DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Science* 258, 1112–1118.

Peliska, J. A., and Benkovic, S. J. (1994). Fidelity of *in vitro* DNA strand transfer reactions catalyzed by HIV-1 reverse transcriptase. *Biochemistry* 33, 3890–3895.

Pilipenko, E. V., Gmyl, A. P., and Agol, V. I. (1995). A model for rearrangements in RNA genomes. *Nucleic Acids Res.* 23, 1870–1875.

Pogany, J., Romero, J., Huang, Q., Sgro, J.-Y., Shang, H., and Bujarski, J. J. (1995). *De novo* generation of defective interfering-like RNAs in broad bean mottle bromovirus. *Virology* 212, 574–586.

Pogany, J. (1997). Molecular studies of the defective RNAs associated with broad bean mottle virus (BBMV). Ph.D. thesis. Northern Illinois University, De Kalb, IL.

Qiao, X., Qiao, J., and Mindich, L. (1997). An *in vitro* system for investigation of heterologous RNA recombination. *Virology* 227, 103–110.

Raffo, A. J., and Dawson, W. O. (1991). Construction of tobacco mosaic virus subgenomic replicons that are replicated and spread systemically in tobacco plants. *Virology* 184, 277–289.

Raju, R., Subramaniam, S. V., and Hajjou, M. (1995). Genesis of Sindbis virus by *in vivo* recombination of nonreplicative RNA precursors. *J. Virol.* 69, 7391–7401.

Romanova, L. I., Blinov, V. M., Tolskaya, E. A., Viktorova, E. G., Kolesnikova, M. S., Guseva, E. A., and Agol, V. I. (1986). The primary structure of crossover regions of intertypic poliovirus recombinants: A model of recombination between RNA genomes. *Virology* 155, 202–213.

Romero, J., Huang, Q., Pogany, J., and Bujarski, J. J. (1993). Characterization of defective interfering RNA components that increases symptom severity of broad bean mottle virus infections. *Virology* 194, 576–584.

Roux, L., Simon, A. E., and Holland, J. J. (1991). Effects of defective interfering viruses on virus replication and pathogenesis *in vitro* and *in vivo. Adv. Virus Res.* 40, 181–211.

Shapira, R., Choi, G. H., Hillman, B. I., and Nuss, D. L. (1991). The contribution of defective RNAs to the complexity of viral-encoded double-stranded RNA populations present in hypovirulent strains of the chestnut blight fungus Cryphonectria parasitica. *EMBO J.* 10, 741–746.

Simon, A. E., and Bujarski, J. J. (1994). RNA–RNA recombination and evolution in virus infected plants. *Annu. Rev. Phytopathol.* 32, 337–362.

Simon, A. E., and Nagy, P. D. (1996). RNA recombination in turnip crinkle virus: Its role in formation of chimeric RNAs, multimers, and in 3′-end repair. *Semin. Virol.* 7, 373–379.

Strauss, J. H., and Strauss, E. G. (1988). Evolution of RNA viruses. *Annu. Rev. Microbiol.* 42, 657–683.

Sullenger, B. A., and Cech, T. R. (1994). Ribozyme-mediated repair of defective mRNA by targeted trans-splicing. *Nature* 371, 619–622.

Triana-Alonso, F. J., Dabrowski, M., Wadzack, J., and Nierhaus, K. H. (1995). Self-coded 3′-extension of run-off transcripts produces aberrant products during *in vitro* transcription with T7 RNA polymerase. *J. Biol. Chem.* 270, 6298–6307.

Tsagris, M., Tabler, M., and Sanger, H. L. (1991). Ribonuclease T1 generates circular RNA molecules from viroid-specific RNA transcripts by cleavage and intramolecular ligation. *Nucleic Acids Res.* 19, 1605–1612.

van der Most, R. G., Heijnen, L., Spaan, W. J. M., and de Groot, R. J. (1992). Homologous RNA recombination allows efficient introduction of site-specific mutations into the genome of coronavirus MHV-A59 via synthetic co-replication RNAs. *Nucleic Acids Res.* 20, 3375–3381.

Wang, D., and Hawley, D. K. (1993). Identification of a 3′–5′ exonuclease activity associated with human RNA polymerase II. *Proc. Natl. Acad. Sci. USA* 90, 843–847.

Wang, H., and Lambowitz, A. M. (1993). The Mauriceville plasmid reverse transcriptase can initiate cDNA synthesis *de novo* and may be related to reverse transcriptase and DNA polymerase progenitor. *Cell* 75, 1071–1081.

White, K. A., and Morris, T. J. (1994a). Nonhomologous RNA recombination in tombusviruses: Generation and evolution of defective interfering RNAs by stepwise deletions. *J. Virol.* 68, 14–24.

White, K. A., and Morris, T. J. (1994b). Recombination between defective tombusvirus RNAs generates functional hybrid genomes. *Proc. Natl. Acad. Sci. USA* 91, 3642–3646.

White, K. A., and Morris, T. J. (1995). RNA determinants of junction site selection in RNA virus determinants and defective interfering RNAs. *RNA* 1, 1029–1040.

Wilson, K. S., and von Hippel, P. (1995). Transcription termination at intrinsic terminators: The role of the RNA hairpin. *Proc. Acad. Natl. Sci. USA* 92, 8793–8797.

Wu, W., Blumberg, B. M., Fay, P. J., and Bambara, R. A. (1995). Strand transfer mediated by Human Immunodeficiency Virus reverse transcriptase *in vitro* is promoted by pausing and results in misincorporation. *J. Biol. Chem.* 270, 325–332.