

## Evolutionary aspects of recombination in RNA viruses

Michael Worobey and Edward C. Holmes

Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

### Introduction

RNA viruses deserve their reputation as Nature's swiftest evolvers. Their high rates of mutation and replication (Holland *et al.*, 1982) allow them to move through sequence space at a pace that often makes their DNA-based hosts' evolution look glacial by comparison. Over the last two decades it has become increasingly clear that many RNA viruses add the capacity to exchange genetic material with one another, and to acquire genes from their hosts, to this evolutionary repertoire. So, in addition to producing prodigious amounts of the raw material of evolution (mutations), these viruses also possess mechanisms that, in principle, allow them both to purge their genomes of accumulated deleterious changes (Muller, 1964) and to create or spread beneficial combinations of mutations in an efficient manner (Fisher, 1930; Muller, 1932), two processes which are not available to clonal organisms.

Two distinct but not mutually exclusive types of genetic exchange operate in RNA viruses. The first, reassortment, occurs only in multipartite viruses and involves swapping one or more of the discrete RNA molecules that make up the segmented viral genome. Antigenic shift in influenza A virus is an example of this sort of genetic exchange and serves as a good illustration of the potential evolutionary significance of such events. A second process, recombination, can occur in either segmented or unsegmented viruses when 'donor' nucleotide sequence is introduced into a single, contiguous 'acceptor' RNA molecule to produce a new RNA containing genetic information from more than one source. In this paper we focus on this type of genetic exchange. First, we briefly review current knowledge of RNA virus recombination and describe new methods for detecting its occurrence using gene sequence data. We then discuss some of the evolutionary implications of virus recombination and some of the constraints that may shape the variety of RNA virus recombination.

### Recombination in RNA viruses

In some cases of RNA virus recombination, the donor sequence neatly replaces a homologous region of the acceptor sequence leaving its structure unchanged. This has been

classified as 'homologous recombination' (Lai, 1992) since it involves not just homologous parental RNAs, but also crossovers at homologous sites. However, this is not always the case; hybrid sequences resulting from aberrant homologous recombination (when similar viruses exchange sequence without maintaining strict alignment) and nonhomologous recombination (recombination between unrelated RNA sequences) are also commonly observed (Lai, 1992).

Despite producing distinct kinds of hybrid RNAs, as well as defective interfering (DI) RNAs (Lazzarini *et al.*, 1981), the different types of recombination appear to be variations on a common theme. To date, almost all studies on the mechanisms of recombination in RNA viruses have supported a copy-choice model, originally proposed in the case of poliovirus (Cooper *et al.*, 1974) and now well studied in a number of experimental systems (Duggal *et al.*, 1997; Jarvis & Kirkegaard, 1992; Kirkegaard & Baltimore, 1986; Nagy & Bujarski, 1995, 1998; Nagy *et al.*, 1998; Simon & Nagy, 1996; for a recent review see Nagy & Simon, 1997). Under this model, hybrid RNAs are formed when the viral RNA-dependent RNA polymerase complex switches, mid-replication, from one RNA molecule to another. This results in homologous recombination if the replicase continues to copy the new strand precisely where it left the old one, and aberrant or nonhomologous recombination if it does not. This template-switching mechanism is fundamentally different from the enzyme-driven breakage-rejoining mechanism of homologous recombination in DNA, not least because it invokes replication as a necessary component of the process. Finally, Chetverin *et al.* (1997) presented evidence for a splicing-like, transesterification mechanism to explain the *in vitro* generation of recombinants between RNAs associated with Q $\beta$  bacteriophage – a possible exception to the copy-choice model of recombination in RNA viruses. Whether such a mechanism operates *in vivo* remains to be seen; however, end-to-end joining is not regarded as a likely mechanism for homologous recombination.

There is now a fairly rich literature documenting recombination in RNA viruses. Many excellent recent reviews have dealt comprehensively with aspects of recombination in experimental and natural settings with respect to animal viruses (Lai, 1992, 1996; Ball, 1997; Strauss & Strauss, 1997), plant viruses (Lai, 1992; Simon & Bujarski, 1994; Roossinck, 1997; Aaziz & Tepfer, 1999) and bacteriophages (Mindich, 1996; Chetverin, 1997). Recently, reports describing hom-

**Author for correspondence:** Michael Worobey.

Fax +44 1865 310447. e-mail michael.worobey@zoo.ox.ac.uk

ologous recombination in rotaviruses (Suzuki *et al.*, 1998) and in hantaviruses (Sibold *et al.*, 1999) have added double-stranded and negative-sense RNA viruses, respectively, to the long list of RNA viruses in which homologous recombination has been detected. The emerging significance of RNA virus recombination is all the more fascinating given the fact that – until the comparatively recent publication by Cooper *et al.* (1974) which showed that mutants of poliovirus could be mapped by recombination analysis – recombination was not thought to be a property of RNA genomes.

### New tools for detecting recombination in viruses

The molecular revolution initiated by the development of PCR has transformed the study of virus recombination. Sequence analysis and phylogenetic techniques have in recent years proven to be extremely effective methods for detecting and characterizing recombination events among RNA viruses both in nature (Gao *et al.*, 1998; Hahn *et al.*, 1988; Holmes *et al.*, 1999; Kusters *et al.*, 1990; Revers *et al.*, 1996; Sibold *et al.*, 1999; Suzuki *et al.*, 1998; Worobey *et al.*, 1999) and in the laboratory (Banner & Lai, 1991; Greene & Allison, 1994; Kotier *et al.*, 1995; Mindich, 1996; Palasingam & Shaklee, 1992; Weiss & Schlesinger, 1991). They offer a way not just to recover information about recombination events that may have occurred long ago or are exceedingly rare (Snijder *et al.*, 1991; Weaver *et al.*, 1997), but also to probe the finest details of the mechanism itself (Banner & Lai, 1991; Olsthoorn & van Duin, 1996). They also provide a means to home in on the precise location of putative crossover points and to test results suggestive of recombination for statistical significance.

Several methods for detecting recombination events and locating breakpoints are graphical in nature, exploiting the fact that many recombinant sequences are mosaics comprising regions with quite different phylogenetic histories. One of these, Split Decomposition analysis (Bandelt & Dress, 1992; Huson, 1998), presents conflicting phylogenetic signal in a single diagram. If no recombination has occurred in the sequences tested, the splits-graph tends to resemble a dichotomously branching phylogenetic tree, because this adequately describes sequence relationships. However, in datasets containing conflicting signal due to the presence of recombinant and hence ‘mosaic’ sequences, the tree-like pattern is often replaced by a more complicated ‘network’ that indicates a history of genetic exchange. It is worth noting that conventional phylogenetics programs are constrained to produce simple branching trees and can lead to serious misinterpretation if sequence alignments are not carefully examined for evidence of recombination prior to tree reconstruction.

Several other graphical applications, including ‘boot-scanning’ (Salminen *et al.*, 1995), ‘PhylPro’ (Weiller, 1998), ‘TOPAL’ (McGuire & Wright, 1998) and ‘DIVERT’ (Gao *et al.*,

1998), use ‘sliding windows’ to detect discordant sequence relationships suggestive of recombination. Bootscanning is an aptly named phylogenetic approach that initially produces a tree from a small window at one end of a sequence alignment and assesses its robustness using bootstrapping. The window is then incrementally shifted along the alignment and a new bootstrap tree is produced for each resulting subset of the alignment. Significant topological changes in the position of a sequence in different windows indicate possible recombination. PhylPro and TOPAL both slide a *pair* of adjacent windows along the sequence alignment. Each of these methods employs a different measure of phylogenetic signal, but in both the phylogenetic information contained in one window is compared to that in the neighbouring window. In the absence of recombination all windows are expected to show similar patterns. On the other hand, if recombination has occurred, some adjacent windows are expected to contain conflicting signal and the difference between them should be greatest when they straddle a recombination breakpoint. DIVERT, the simplest of the sliding window graphical methods (and often the most effective), outputs a graph of genetic distance comparisons between a chosen sequence and comparison sequences, which can show runs of sequence similarity and dissimilarity suggestive of recombination. Diversity plots have been used to great effect in the search for recombinant human immunodeficiency virus and simian immunodeficiency virus strains (Gao *et al.*, 1998, 1999) and are ideally suited for detection of RNA virus homologous recombination (Worobey *et al.*, 1999).

Many of these programs permit a simple qualitative assessment of possible recombination breakpoints based on the visual analysis of their output. However, for cases where putative recombinants and reasonably close relatives of their acceptor and donor sequences are available, more sophisticated procedures exist for locating crossover points. Informative Sites Analysis (Robertson *et al.*, 1995), a parsimony-based adaptation of the maximum  $\chi^2$  test (Maynard Smith, 1992), uses the distribution of polymorphic sites between a probable recombinant and its putative ‘parents’ to estimate recombination junctions. The results can then be compared to randomized distributions of polymorphic sites to assess their significance. A similar method, LARD (Holmes *et al.*, 1999), uses a maximum likelihood method to infer the optimal breakpoints for a possible recombinant, then uses simulated sequences to test the statistical significance of the results.

Some other methods attempt to quantify the amount of recombination between a set of sequences, rather than document specific recombination events, often using the degree of linkage equilibrium. One way this can be done is with the Index of Association. Using this statistical test, which was designed to detect associations between alleles at different loci, it is possible to measure the extent of linkage equilibrium within populations (Maynard Smith *et al.*, 1993). Another, based on a direct phylogenetic analysis, is the Homoplasmy Test

(Maynard Smith & Smith, 1998). Here, the number of homoplastic (i.e. convergent and parallel) base changes in data observed after construction of a maximum parsimony tree is compared to that number expected by chance. Excessive homoplasies are the fingerprint of recombination. (WWW links to the Homoplasy Test package and to TOPAL, PhylPro, bootscanning and the maximum  $\chi^2$  method can be found at [http://igs-server.cnrs-mrs.fr/anrs/phylogenetics/RAP\\_links.html](http://igs-server.cnrs-mrs.fr/anrs/phylogenetics/RAP_links.html).)

## How clonal are viruses?

The impressive tally of recombinant ‘hopeful monsters’ – dramatically altered hybrid viruses that have passed the test of natural selection and emerged as successful new viruses – attests to the power of recombination as an evolutionary force in RNA viruses (Allison *et al.*, 1989; Gibbs & Cooper, 1995; Herrewegh *et al.*, 1998; Luytjes *et al.*, 1988; Revers *et al.*, 1996; Röhm *et al.*, 1996; Snijder *et al.*, 1991; Suzuki *et al.*, 1998; Weaver *et al.*, 1997). In some cases, evolutionary evidence suggests that newly formed recombinant strains may have initially exhibited decreased functionality, but subsequently evolved to compensate for these effects. For example, the Western equine encephalitis (WEE) complex viruses are the product of a single recombination event that occurred between Eastern equine encephalitis virus (EEEV) and a Sindbis-like virus, probably within the last 2000 years (Weaver *et al.*, 1997). Sequence analysis of WEEV indicated that, subsequent to the recombination event, its EEEV-like capsid protein evolved to become more like a Sindbis-virus capsid, possibly because it needs to interact with Sindbis-like glycoproteins during virus budding (Hahn *et al.*, 1988). Conversely, in a sort of evolutionary compromise, almost all the amino acid changes in the Sindbis-like glycoproteins have been to residues that match those of EEEV. It is possible that the new antigenic properties conferred by the Sindbis-like glycoproteins of this predominantly EEEV-like hybrid were sufficiently advantageous to off-set what appears to have been a significant mismatch between its recombinant structural proteins.

Whatever the circumstances of the survival and subsequent diversification of particular recombinants, it is now evident that many pass through the narrow gates of natural selection and contribute to the diversity seen in RNA viruses. The production of new strains having genomes comprising regions with different histories has important implications for the way we think about virus evolution. For one, it means that there is no single phylogenetic tree that can describe the evolutionary relationships between viruses; the recombinative nature of viruses simply precludes the possibility of a ‘true’ phylogenetic taxonomy. Far from being an incidental process best ignored when considering virus relationships and history, recombination may have played a crucial role in generating many of the taxonomic groups we recognize (Koonin & Dolja, 1993) –

today’s hopeful monster giving rise to tomorrow’s genus or family of viruses. Some studies of recombination have even led to recommendations that higher-than-family taxonomic units should be avoided altogether (Goldbach, 1992).

The inappropriateness of tree-like representations of evolutionary relatedness may also apply within some virus species. In principle, the possibility of homologous recombination between similar strains means that the population structure in viruses could range from completely clonal, if no recombination has taken place, to panmictic, if recombination has been common enough to effectively randomize loci. Although significant progress has been made studying population structure in other organisms (Maynard Smith *et al.*, 1993), similar work remains to be done for RNA viruses.

Many taxonomic groupings of viruses – for example both the *Togaviridae* and the *Coronaviridae* – include members that, while clearly sharing homologous genes, differ in the order in which these genes are organized along the genome, another indication that recombination has been important in their evolutionary history. Other cases – exemplified by the haemagglutinin-esterase gene known to be present in at least three virus genera (Snijder *et al.*, 1991) – show recombination as the engine driving modular evolution, whereby functional modules from different sources are brought together to create new viruses. The discovery of recombination in an increasing number of viruses, in addition to presenting phylogenetic and taxonomic difficulties, challenges the desirability of using short sequence regions as markers for entire virus genomes (for instance in molecular epidemiology studies) since they may not accurately reflect true genetic or antigenic characteristics.

## Evolutionary advantages of recombination

Theoretical explanations for the evolution of recombination and hence some aspects of sexual reproduction tend to fit into one of two standard classes: (1) that it enables the creation and spread of advantageous traits, and (2) that it permits the removal of deleterious genes (for a review see Hurst & Peck, 1996). This second explanation is often linked to the notion of ‘Muller’s ratchet’ – the random loss of those individuals in a population having the fewest deleterious alleles (Muller, 1964). Muller’s ratchet predicts the gradual build-up of deleterious alleles in a finite, asexual population. Experiments with RNA viruses, one of only a handful of organisms in which this hypothesis has actually been put to the test, have generally supported its operation and shown decreased fitness for populations in which it occurs (Chao *et al.*, 1992). Experimental evidence (Chao *et al.*, 1997) also shows that sex (in this case reassortment) can reduce the mutational load in a population and so help it escape from accumulated deleterious effects.

Although such direct experimental evidence has yet to demonstrate a similar advantage for recombination, in principle it too could serve to efficiently remove disadvantageous alleles from a population by combining mutation-free parts of

different genomes. Indeed, suggestions have been made that reassortment in segmented RNA viruses and recombination in monopartite RNA viruses represent alternative evolutionary strategies for genetic exchange in this group (Chao *et al.*, 1992). While this idea is fascinating, it is interesting to note that reassortment and recombination are not mutually exclusive and that several segmented viruses also experience recombination, sometimes frequently. These include the bacteriophage  $\phi 6$  (Mindich *et al.*, 1992), rotaviruses (Suzuki *et al.*, 1998), influenza A virus (Khatchikian *et al.*, 1989), hantaviruses (Sibold *et al.*, 1999), flock house virus (Li & Ball, 1993) and many plant viruses (Bujarski & Kaesberg, 1986; Greene & Allison, 1994; Robinson, 1994; Rott *et al.*, 1991). As if to prove this point, Masuta *et al.* (1998) recently reported an interspecific hybrid of two cucumoviruses that arose by both reassortment and recombination.

Nonetheless, a great deal of evidence indicates that some RNA viruses do benefit from the genome-purging effects of recombination. A multitude of experimental studies have shown that weak or even non-replicative mutant strains can recombine to form viable, highly fit viruses. Examples include the functional chimeras formed between nonreplicating RNAs and DI RNAs of tombusviruses (White & Morris, 1994), infectious recombinants produced by different combinations of mutationally altered Sindbis virus RNAs (Raju *et al.*, 1995; Weiss & Schlesinger, 1991) and wild-type revertant recombinants of Q $\beta$  phage mutants (Palasingam & Shaklee, 1992) and of bromovirus mutants (Rao & Hall, 1993).

Plant viruses have also been observed to repair their genomes by recombining with host transgene transcripts (Borja *et al.*, 1999; Gal-On *et al.*, 1998; Greene & Allison, 1994; Rubio *et al.*, 1999). Similarly, in one experiment with a deletion mutant of mouse hepatitis virus (MHV) transfected with a synthetic RNA that contained the deleted region (Koetzner *et al.*, 1992), and another with an influenza A virus mutant with a damaged neuraminidase gene (Bergmann *et al.*, 1992), recombination successfully repaired defective genes. Studies of recombination in bacteriophages, too, indicate a repair function for recombination (Mindich *et al.*, 1994). RNA recombination even appears to provide a telomerase-like function by repairing the 3' ends of satellite RNAs of both turnip crinkle virus (Burgyan & García-Arenal, 1998) and cucumber mosaic virus (Simon & Nagy, 1996).

Unintentional 'natural experiments' with some viruses point to the same conclusion. The frequent recovery of recombinant isolates of poliovirus (Georgescu *et al.*, 1994; Kew & Nottay, 1984) and infectious bronchitis virus (Jia *et al.*, 1995; Kusters *et al.*, 1990; Wang *et al.*, 1994) that result from recombination involving vaccine strains shows that recombination has the potential to produce 'escape mutants' in nature as well as in experiments. Recently, recombination has also been detected in other RNA viruses for which multivalent vaccines are in use or in trials (Holmes *et al.*, 1999; Suzuki *et al.*, 1998; Worobey *et al.*, 1999). We think the potential for

recombination to produce new pathogenic hybrid strains, and the possible impact of such escape recombination, needs to be carefully considered whenever multivalent live-attenuated vaccines are used to control RNA viruses. Assumptions that recombination either does not happen or is unimportant in RNA viruses have a history of being proved wrong.

In addition to the evidence favouring a role for genetic exchange in eliminating deleterious alleles, many recombinant RNA virus strains provide ample indication that recombination can generate beneficial new variation. In some viruses this new variation is achieved by borrowing genetic material from their hosts. One intriguing example of this is bovine viral diarrhoea virus (BVDV), a pestivirus that recombines with host cellular protein-coding RNA. As a result of virus–host recombinations, cytopathogenic BVDVs can develop from non-cytopathogenic ones and cause a lethal syndrome, mucosal disease, in the hosts (Meyers *et al.*, 1989). Influenza A virus has also been observed to recombine with cellular RNA, resulting in increased pathogenicity for the hybrid viruses (Khatchikian *et al.*, 1989). Recombination between virus and host genetic material evidently occurs in plant viruses as well, as illustrated by a luteovirus isolate with 5'-terminal sequence derived from a chloroplast exon (Mayo & Jolly, 1991) and closteroviruses which have acquired host cellular protein-coding genes (Dolja *et al.*, 1994) which are nonessential for replication and virion production (Peremyslov *et al.*, 1998).

A link between recombination and increased pathogenicity has also been revealed in cases that do not involve recombination with host genes. Template jumping during replication in viruses infecting cats has produced, on multiple occasions, the pathogenic strains known as feline infectious peritonitis viruses (FIPVs) by altering asymptomatic feline enteric coronaviruses, differing from them only by deletions of around 100 bp in predictable locations (Vennema *et al.*, 1998). Another coronavirus, feline coronavirus (FCoV) type II, appears to be a homologous (or aberrant homologous) recombinant of FCoV type I and canine coronavirus (Herrewegh *et al.*, 1995). Like FIPVs, FCoV type II viruses may have arisen on different occasions from separate recombination events (Motokawa *et al.*, 1996).

Experimental studies provide further signs of the ability of recombination to generate useful, new variation. In one particularly striking display of this, MS2 phage mutants lacking the sequence for important stem-and-loop secondary structures repeatedly reconstructed them via nonhomologous recombination (Olsthoorn & van Duin, 1996).

## Constraints on recombination

Recombination clearly plays a significant role in the evolution of RNA viruses by generating genetic variation, by reducing mutational load, and by producing new viruses. We expect that with current advances in sequencing and sequence analysis many more examples of hybrid viruses, produced by

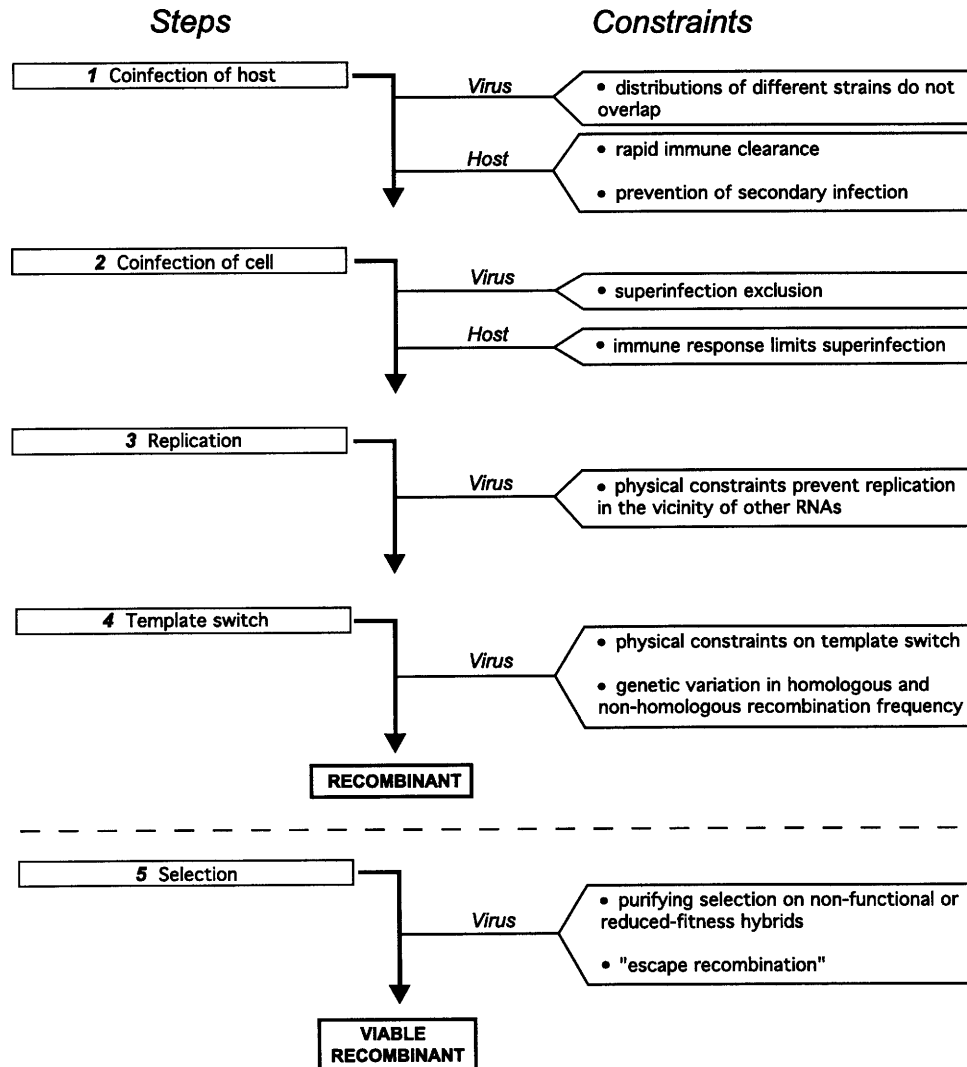


Fig. 1. A model of the production of a viable recombinant RNA virus by copy-choice, and possible constraints against this process. Steps 1–4 are necessary to produce any recombinant RNA, while step 5 must be passed to generate a competent, replicating hybrid virus. Virus or host constraints could block recombination by limiting or preventing the completion of any of the five steps.

recombination between different strains, will soon be found. However, it is clearly not the case that all RNA viruses are equally prone to recombination. It has still not been detected in several viruses despite strenuous searches (e.g. Bilsel *et al.*, 1990), although some – not otherwise known to recombine – nevertheless produce DI RNAs. Amongst those known to produce hybrids, the frequency of recombinants detected in natural or experimental studies varies markedly. Given the potential advantages of recombination, why is there apparently so much variation between viruses in its occurrence? While it is too soon to provide a definitive answer to this crucial question, it is possible to dissect the process that gives rise to recombinants and to consider the constraints that could act at each stage to inhibit it. A simple model of recombination between different RNA viruses, with possible constraints, is presented in Fig. 1.

In a sense, recombinogenic viruses are all alike in that they successfully pass through each stage outlined in the model. Every non-recombining virus, on the other hand, is different in its own way since constraints that block recombination could act by breaking *any* link in the chain and could involve not just viral genetic factors, but host and ecological factors particular to that virus. The first prerequisite for successful recombination (Fig. 1) is that an individual host must be infected by different virus strains. [This is not quite true, of course, since (1) recombination sometimes involves host RNA, (2) recombination could occur between viruses that have diverged within a clonally infected individual, and (3) evolutionarily invisible recombination could occur between identical RNA molecules.] Host coinfection might never occur with some viruses simply because their divergent forms do not usually overlap in space and time. Multivalent live-attenuated vaccines

can be seen as potential risks in this context since they could effectively release some viruses from this constraint. Host factors could act at this stage too if, for instance, an immune response reduces the window for simultaneous infection by quickly clearing a virus, or prevents superinfection altogether by blocking secondary infections.

Having successfully coinfecting a single host, divergent viruses must next coinfect a single cell if recombination is to proceed. This step could be blocked by host factors, either by an immune response that keeps virus numbers low enough to prevent multiple infection of any individual cell, or by host cell genetic factors that block entry of more than one virus particle into a cell (Danis *et al.*, 1993). Viral factors, interestingly, might also enforce significant constraints at this stage of the model. Recent evidence demonstrates that intracellular competition can be costly to viruses that infect the same host cell (Turner & Chao, 1998). Those that can selfishly keep cells to themselves by limiting or preventing coinfection should be selectively favoured, and many have evolved mechanisms to do just this (Simon *et al.*, 1990; Singh *et al.*, 1997; Turner *et al.*, 1999). One of these, vesicular stomatitis virus (VSV), is an RNA virus in which recombination between different strains has not been detected. Might superinfection exclusion be a constraint on recombination in VSV? Another, the segmented bacteriophage  $\phi 6$ , seems to limit excessive superinfection but not to the extreme of one-virus-per-cell that would preclude genetic exchange. Instead, it appears to have evolved an optimal coinfection limit of two to three viruses per cell, presumably to balance the costs of intracellular competition with the benefits of reassortment (Turner *et al.*, 1999). Since the advantage (and cost) of recombination in any particular virus will be mediated by such factors as the selective pressure for novel variation, the importance of interactions between different parts of the genome, as well as the virus mutation rate and population size, we should expect different optima (and therefore different degrees of constraint) in different cases.

If divergent viruses manage to infect the same cell, the next step is simply for one of them to replicate in the presence of the RNA of the other. This is not necessarily inevitable even in coinfecting cells. The replication of the  $\phi 6$  RNAs, for example, takes place within a procapsid and it is thought that the entry of two different RNA molecules of the same genomic segment into this sequestered environment is impossible or at least very rare (Mindich *et al.*, 1992). This could explain the lack of homologous recombination in this phage. Thus the vagaries of RNA replication in certain viruses could impose physical constraints on the production of hybrids.

Template switching by the viral replicase, the mechanism whereby recombinant RNA molecules are actually created, may also be limited by physical constraints. The negative-strand RNA viruses, for example, whose genomes are packaged into filamentous ribonucleoprotein structures by association with N protein, may be less permissive than other RNA viruses to copy-choice recombination. And perhaps the most im-

portant physical constraint on template switching – particularly with respect to homologous recombination – is simply the extent of sequence dissimilarity between potentially recombining genomes. Finally, genetic variation in the susceptibility of the viral replicase to jumping (Bujarski & Nagy, 1996) no doubt plays a central role in determining how often and by what mechanism particular viruses recombine.

Recombination occurs when these first four steps are fulfilled. Whether incipient recombinants persist, however, depends on the fifth and final step, the selective separation of the wheat from the chaff among hybrids. Although there is strong evidence that genetic exchange can offer advantages in some circumstances, random recombination no doubt destroys more good alleles than it creates. PCR studies which have made possible the characterization of the initial products of recombination – those present prior to removal by selection (Banner & Lai, 1991; Desport *et al.*, 1998; Jarvis & Kirkegaard, 1992) – have produced important insights. Banner and Lai's study of coronaviruses (1991), for example, showed that the initial recombination events in their MHV system were almost entirely randomly distributed along the sequence investigated. It was only after passage through cell culture, with the opportunity for selection to remove less fit variants, that crossover sites became 'localized' to just a small area of the region examined. With enough passages, the recombinants disappeared altogether. These results indicated that 'recombination hotspots' can actually be the result of natural selection on a pool of random recombination crossover junctions, as opposed to elevated recombination rates in particular regions. Crucially, they also suggested that recombination may be more common than often assumed, but may go undetected because of the action of strong purifying selection which will remove new, deleterious combinations of mutations. In light of these studies it is clear that what is meant by 'recombination frequency' – a term usually used without specified units – depends critically on whether we are assessing recombination events before or after selection has acted. A virus which often produces hybrid RNAs under laboratory conditions may very rarely – or even never – be found to recombine in nature. This difference is analogous to the important distinction between the rate of mutation and the rate of substitution.

Negative selection against non-functional hybrids or those with decreased fitness may impose the strongest constraints of all on the appearance of recombinants. In viruses for which the evolutionary costs of recombination outweigh the benefits, though they may be mechanistically capable of genetic exchange, strong selection will guarantee the elimination of hybrid genomes.

## Conclusion

Determining the constraints that operate on recombination offers a promising path to a fuller understanding of its importance in the evolution of RNA viruses. However, outlines

of the big picture are already clearly visible. It seems certain that genetic exchange plays a key role in several virus groups and that it has shaped a good deal of the diversity – both ancient and recent – that exists in them. Thus, evolutionary knowledge about recombination impacts on many aspects of the study of RNA viruses, from the broadest investigations of virus taxonomy, to the finest details of molecular epidemiology and vaccine design. A flood of viral gene sequence data and the availability of new and powerful phylogenetic methods is making the detection and characterization of recombination ever easier, and the list of viruses showing recombination continues to grow. The evidence for recombination, not only between closely related viruses but also among distantly related viruses, positive-sense and negative-sense viruses, DI RNAs and viruses, satellite RNAs and viruses, and even with host RNAs, suggests that almost any genetic material can be grist for the polymerase's mill. Of all the tricks up the viral evolutionary sleeve, surely recombination is one of the most deft.

## References

- Aaziz, R. & Tepfer, M. (1999).** Recombination in RNA viruses and in virus-resistant transgenic plants. *Journal of General Virology* **80**, 1339–1346.
- Allison, R. F., Janda, M. & Ahlquist, P. (1989).** Sequence of cowpea chlorotic mottle virus RNA 2 and RNA 3 and evidence of a recombination event during bromovirus evolution. *Virology* **172**, 321–330.
- Ball, L. A. (1997).** Nodavirus RNA recombination. *Seminars in Virology* **8**, 95–100.
- Bandelt, H. J. & Dress, A. W. M. (1992).** Split Decomposition: a new and useful approach for the phylogenetic analysis of distance data. *Molecular Phylogenetics and Evolution* **1**, 242–252.
- Banner, L. R. & Lai, M. M. C. (1991).** Random nature of coronavirus RNA recombination in the absence of selection pressure. *Virology* **185**, 441–445.
- Bergmann, M., García-Sastre, A. & Palese, P. (1992).** Transfection-mediated recombination of influenza A virus. *Journal of Virology* **66**, 7576–7580.
- Bilsel, P. A., Rowe, J. E., Fitch, W. M. & Nichol, S. T. (1990).** Phosphoprotein and nucleocapsid protein evolution of vesicular stomatitis virus New Jersey. *Journal of Virology* **64**, 2498–2504.
- Borja, M., Rubio, T., Scholthof, H. B. & Jackson, A. O. (1999).** Restoration of wild-type virus by double recombination of tomosvirus mutants with a host transgene. *Molecular Plant–Microbe Interactions* **12**, 153–162.
- Bujarski, J. J. & Kaesberg, P. (1986).** Genetic recombination between RNA components of a multipartite plant virus. *Nature* **321**, 528–531.
- Bujarski, J. J. & Nagy, P. D. (1996).** Different mechanisms of homologous and nonhomologous recombination in brome mosaic virus: role of RNA sequences and replicase proteins. *Seminars in Virology* **7**, 363–372.
- Burgyan, J. & García-Arenal, F. (1998).** Template-independent repair of the 3' end of cucumber mosaic virus satellite RNA controlled by RNAs 1 and 2 of helper virus. *Journal of Virology* **72**, 5061–5066.
- Chao, L., Tran, T. R. & Matthews, C. (1992).** Muller's ratchet and the advantage of sex in the RNA virus  $\phi 6$ . *Evolution* **46**, 289–299.
- Chao, L., Tran, T. T. & Tran, T. T. (1997).** The advantage of sex in the RNA virus  $\phi 6$ . *Genetics* **147**, 953–959.
- Chetverin, A. B. (1997).** Recombination in bacteriophage Q $\beta$  its satellite RNAs: the *in vivo* and *in vitro* studies. *Seminars in Virology* **8**, 121–129.
- Chetverin, A. B., Chetverina, H. V., Demidenko, A. A. & Ugarov, V. I. (1997).** Nonhomologous RNA recombination in a cell-free system: evidence for a transesterification mechanism guided by secondary structure. *Cell* **88**, 503–513.
- Cooper, P. D., Steiner-Pryor, A., Scotti, P. D. & DeLong, D. (1974).** On the nature of poliovirus genetic recombinants. *Journal of General Virology* **23**, 41–49.
- Danis, C., Mabrouk, T., Garzon, S. & Lemay, G. (1993).** Establishment of persistent reovirus infection in SC1 cells: absence of protein-synthesis inhibition and increased level of double-stranded RNA-activated protein-kinase. *Virus Research* **27**, 253–265.
- Desport, M., Collins, M. E. & Brownlie, J. (1998).** Genome instability in BVDV: an examination of the sequence and structural influences on RNA recombination. *Virology* **246**, 352–361.
- Dolja, V. V., Karasev, A. V. & Koonin, E. V. (1994).** Molecular biology and evolution of closteroviruses: sophisticated buildup of large RNA genomes. *Annual Review of Phytopathology* **32**, 261–285.
- Duggal, R., Cuconati, A., Gromeier, M. & Wimmer, E. (1997).** Genetic recombination of poliovirus in a cell-free system. *Proceedings of the National Academy of Sciences, USA* **94**, 13786–13791.
- Fisher, R. A. (1930).** *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.
- Gal-On, A., Meiri, E., Raccach, B. & Gaba, V. (1998).** Recombination of engineered defective RNA species produces infective potyvirus in planta. *Journal of Virology* **72**, 5268–5270.
- Gao, F., Robertson, D. L., Carruthers, C. D., Morrison, S. G., Jian, B. X., Chen, Y. L., Barré-Sinoussi, F., Girard, M., Srinivasan, A., Abimiku, A. G., Shaw, G. M., Sharp, P. M. & Hahn, B. H. (1998).** A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1. *Journal of Virology* **72**, 5680–5698.
- Gao, F., Bailes, E., Robertson, D. L., Chen, Y. L., Rodenburg, C. M., Michael, S. F., Cummins, L. B., Arthur, L. O., Peeters, M., Shaw, G. M., Sharp, P. M. & Hahn, B. H. (1999).** Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**, 436–441.
- Georgescu, M. M., Delpeyroux, F., Tardypanit, M., Balanant, J., Combiescu, M., Combiescu, A. A., Guillot, S. & Crainic, R. (1994).** High diversity of poliovirus strains isolated from the central nervous system from patients with vaccine-associated paralytic poliomyelitis. *Journal of Virology* **68**, 8089–8101.
- Gibbs, M. J. & Cooper, J. I. (1995).** A recombinational event in the history of luteoviruses probably induced by base-pairing between the genomes of 2 distinct viruses. *Virology* **206**, 1129–1132.
- Goldbach, R. (1992).** The recombinative nature of potyviruses: implications for setting up a true phylogenetic taxonomy. *Archives of Virology* **5** (Suppl.), 299–304.
- Greene, A. E. & Allison, R. F. (1994).** Recombination between viral RNA and transgenic plant transcripts. *Science* **263**, 1423–1425.
- Hahn, C. S., Lustig, S., Strauss, E. G. & Strauss, J. H. (1988).** Western equine encephalitis virus is a recombinant virus. *Proceedings of the National Academy of Sciences, USA* **85**, 5997–6001.
- Herrewegh, A. A. P. M., Vennema, H., Horzinek, M. C., Rottier, P. J. M. & de Groot, R. J. (1995).** The molecular genetics of feline coronaviruses: comparative sequence analysis of the ORF7a/7b transcription unit of different biotypes. *Virology* **212**, 622–631.

- Herrewegh, A. A. P. M., Smeenk, I., Horzinek, M. C., Rottier, P. J. M. & de Groot, R. J. (1998). Feline coronavirus type II strains 79-1683 and 79-1146 originate from a double recombination between feline coronavirus type I and canine coronavirus. *Journal of Virology* **72**, 4508–4514.
- Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S. & Vandepol, S. (1982). Rapid evolution of RNA genomes. *Science* **215**, 1577–1585.
- Holmes, E. C., Worobey, M. & Rambaut, A. (1999). Phylogenetic evidence for recombination in dengue virus. *Molecular Biology and Evolution* **16**, 405–409.
- Hurst, L. D. & Peck, J. R. (1996). Recent advances in understanding of the evolution and maintenance of sex. *Trends in Ecology & Evolution* **11**, 46–52.
- Huson, D. H. (1998). SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**, 68–73.
- Jarvis, T. C. & Kirkegaard, K. (1992). Poliovirus RNA recombination: mechanistic studies in the absence of selection. *EMBO Journal* **11**, 3135–3145.
- Jia, W., Karaca, K., Parrish, C. R. & Naqi, S. A. (1995). A novel variant of avian infectious bronchitis virus resulting from recombination among 3 different strains. *Archives of Virology* **140**, 259–271.
- Kew, O. M. & Nottay, B. K. (1984). Evolution of the oral polio vaccine strains in humans occurs by both mutation and intramolecular recombination. In *Modern Approaches to Vaccines*, pp. 357–363. Edited by R. M. Chanock & R. A. Lerner. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Khatchikian, D., Orlich, M. & Rott, R. (1989). Increased viral pathogenicity after insertion of a 28S ribosomal-RNA sequence into the hemagglutinin gene of an influenza virus. *Nature* **340**, 156–157.
- Kirkegaard, K. & 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. & Masters, P. S. (1992). Repair and mutagenesis of the genome of a deletion mutant of the coronavirus mouse hepatitis virus by targeted RNA recombination. *Journal of Virology* **66**, 1841–1848.
- Koonin, E. V. & Dolja, V. V. (1993). Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino-acid sequences. *Critical Reviews in Biochemistry and Molecular Biology* **28**, 375–430.
- Kotier, S. A., Cavanagh, D. & Britton, P. (1995). Experimental evidence of recombination in coronavirus infectious bronchitis virus. *Virology* **213**, 569–580.
- Kusters, J. G., Jager, E. J., Niesters, H. G. M. & Vanderzeijst, B. A. M. (1990). Sequence evidence for RNA recombination in field isolates of avian coronavirus infectious bronchitis virus. *Vaccine* **8**, 605–608.
- Lai, M. M. C. (1992). RNA recombination in animal and plant viruses. *Microbiological Reviews* **56**, 61–79.
- Lai, M. M. C. (1996). Recombination in large RNA viruses: coronaviruses. *Seminars in Virology* **7**, 381–388.
- Lazzarini, R. A., Keene, J. D. & Schubert, M. (1981). The origins of defective interfering particles of the negative-strand RNA viruses. *Cell* **26**, 145–154.
- Li, Y. & Ball, L. A. (1993). Nonhomologous RNA recombination during negative-strand synthesis of flock house virus RNA. *Journal of Virology* **67**, 3854–3860.
- Luytjes, W., Bredenbeek, P. J., Noten, A. F. H., Horzinek, M. C. & Spaan, W. J. M. (1988). Sequence of mouse hepatitis virus A59 messenger RNA 2: indications for RNA recombination between coronaviruses and influenza C virus. *Virology* **166**, 415–422.
- McGuire, G. & Wright, F. (1998). TOPAL: recombination detection in DNA and protein sequences. *Bioinformatics* **14**, 219–220.
- Masuta, C., Ueda, S., Suzuki, M. & Uyeda, I. (1998). Evolution of a quadripartite hybrid virus by interspecific exchange and recombination between replicase components of two related tripartite RNA viruses. *Proceedings of the National Academy of Sciences, USA* **95**, 10487–10492.
- Maynard Smith, J. (1992). Analyzing the mosaic structure of genes. *Journal of Molecular Evolution* **34**, 126–129.
- Maynard Smith, J. & Smith, N. H. (1998). Detecting recombination from gene trees. *Molecular Biology and Evolution* **15**, 590–599.
- Maynard Smith, J., Smith, N. H., O'Rourke, M. & Spratt, B. G. (1993). How clonal are bacteria? *Proceedings of the National Academy of Sciences, USA* **90**, 4384–4388.
- Mayo, M. A. & Jolly, C. A. (1991). The 5'-terminal sequence of potato leafroll virus RNA: evidence of recombination between virus and host RNA. *Journal of General Virology* **72**, 2591–2595.
- Meyers, G., Rumenapf, T. & Thiel, H. J. (1989). Ubiquitin in a togavirus. *Nature* **341**, 491.
- Mindich, L. (1996). Heterologous recombination in the segmented dsRNA genome of bacteriophage  $\phi 6$ . *Seminars in Virology* **7**, 389–397.
- Mindich, L., Qiao, X. Y., Onodera, S., Gottlieb, P. & Strassman, J. (1992). Heterologous recombination in the double-stranded RNA bacteriophage  $\phi 6$ . *Journal of Virology* **66**, 2605–2610.
- Mindich, L., Qiao, X. Y., Onodera, S., Gottlieb, P. & Frilander, M. (1994). RNA structural requirements for stability and minus-strand synthesis in the dsRNA bacteriophage  $\phi 6$ . *Virology* **202**, 258–263.
- Motokawa, K., Hohdatsu, T., Hashimoto, H. & Koyama, H. (1996). Comparison of the amino acid sequence and phylogenetic analysis of the peplomer, integral membrane and nucleocapsid proteins of feline, canine and porcine coronaviruses. *Microbiology & Immunology* **40**, 425–433.
- Muller, H. J. (1932). Some genetic aspects of sex. *American Naturalist* **66**, 118–138.
- Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research* **1**, 2–9.
- Nagy, P. D. & Bujarski, J. J. (1995). Efficient system of homologous RNA recombination in brome mosaic virus: sequence and structure requirements and accuracy of crossovers. *Journal of Virology* **69**, 131–140.
- Nagy, P. D. & Bujarski, J. J. (1998). Silencing homologous RNA recombination hot spots with GC-rich sequences in brome mosaic virus. *Journal of Virology* **72**, 1122–1130.
- Nagy, P. D. & Simon, A. E. (1997). New insights into the mechanisms of RNA recombination. *Virology* **235**, 1–9.
- Nagy, P. D., Zhang, C. X. & Simon, A. E. (1998). Dissecting RNA recombination in vitro: role of RNA sequences and the viral replicase. *EMBO Journal* **17**, 2392–2403.
- Olsthoorn, R. C. L. & van Duin, J. (1996). Evolutionary reconstruction of a hairpin deleted from the genome of an RNA virus. *Proceedings of the National Academy of Sciences, USA* **93**, 12256–12261.
- Palasingam, K. & Shaklee, P. N. (1992). Reversion of Q $\beta$  RNA phage mutants by homologous RNA recombination. *Journal of Virology* **66**, 2435–2442.
- Peremyslov, V. V., Hagiwara, Y. & Dolja, V. V. (1998). Genes required for replication of the 15.5-kilobase RNA genome of a plant closterovirus. *Journal of Virology* **72**, 5870–5876.
- Raju, R., Subramaniam, S. V. & Hajjou, M. (1995). Genesis of Sindbis virus by in vivo recombination of nonreplicative RNA precursors. *Journal of Virology* **69**, 7391–7401.
- Rao, A. L. N. & Hall, T. C. (1993). Recombination and polymerase error

- facilitate restoration of infectivity in bromo mosaic virus. *Journal of Virology* **67**, 969–979.
- Revers, F., Le Gall, O., Candresse, T., Le Romancer, M. & Dunez, J. (1996).** Frequent occurrence of recombinant potyvirus isolates. *Journal of General Virology* **77**, 1953–1965.
- Robertson, D. L., Hahn, B. H. & Sharp, P. M. (1995).** Recombination in AIDS viruses. *Journal of Molecular Evolution* **40**, 249–259.
- Robinson, D. J. (1994).** Sequences at the ends of RNA 2 of I6, a recombinant tobnavirus. *Archives of Virology*, 245–251.
- Röhm, C., Zhou, N. A., Suss, J. C., Mackenzie, J. & Webster, R. G. (1996).** Characterization of a novel influenza hemagglutinin, H15: criteria for determination of influenza A subtypes. *Virology* **217**, 508–516.
- Roossinck, M. J. (1997).** Mechanisms of plant virus evolution. *Annual Review of Phytopathology* **35**, 191–209.
- Rott, M. E., Tremaine, J. H. & Rochon, D. M. (1991).** Comparison of the 5' and 3' termini of tomato ringspot virus RNA 1 and RNA 2: evidence for RNA recombination. *Virology* **185**, 468–472.
- Rubio, T., Borja, M., Scholthof, H. B. & Jackson, A. O. (1999).** Recombination with host transgenes and effects on virus evolution: an overview and opinion. *Molecular Plant–Microbe Interactions* **12**, 87–92.
- Salminen, M. O., Carr, J. K., Burke, D. S. & McCutchan, F. E. (1995).** Identification of breakpoints in intergenotypic recombinants of HIV type-1 by bootscanning. *AIDS Research and Human Retroviruses* **11**, 1423–1425.
- Sibold, C., Meisel, H., Kruger, D. H., Labuda, M., Lysy, J., Kozuch, O., Pejcoch, M., Vaheri, A. & Plyusnin, A. (1999).** Recombination in Tula hantavirus evolution: analysis of genetic lineages from Slovakia. *Journal of Virology* **73**, 667–675.
- Simon, A. E. & Bujarski, J. J. (1994).** RNA–RNA recombination and evolution in virus-infected plants. *Annual Review of Phytopathology* **32**, 337–362.
- Simon, A. E. & Nagy, P. D. (1996).** RNA recombination in turnip crinkle virus: its role in formation of chimeric RNAs, multimers, and in 3'-end repair. *Seminars in Virology* **7**, 373–379.
- Simon, K. O., Cardamone, J. J., Whitaker-Dowling, P. A., Youngner, J. & Widnell, C. C. (1990).** Cellular mechanisms in the superinfection exclusion of vesicular stomatitis virus. *Virology* **177**, 375–379.
- Singh, I., Suomalainen, M., Varadarajan, S., Garoff, H. & Helenius, A. (1997).** Multiple mechanisms of entry and uncoating of superinfecting Semliki Forest virus. *Virology* **231**, 59–71.
- Snijder, E. J., Denboon, J. A., Horzinek, M. C. & Spaan, W. J. M. (1991).** Comparison of the genome organization of toroviruses and coronaviruses: evidence for 2 nonhomologous RNA recombination events during Berne virus evolution. *Virology* **180**, 448–452.
- Strauss, J. H. & Strauss, E. G. (1997).** Recombination in Alphaviruses. *Seminars in Virology* **8**, 85–94.
- Suzuki, Y., Gojobori, T. & Nakagomi, O. (1998).** Intragenic recombinations in rotaviruses. *FEBS Letters* **427**, 183–187.
- Turner, P. E. & Chao, L. (1998).** Sex and the evolution of intrahost competition in RNA virus  $\phi$ 6. *Genetics* **150**, 523–532.
- Turner, P. E., Burch, C. L., Hanley, K. A. & Chao, L. (1999).** Hybrid frequencies confirm limit to coinfection in the RNA bacteriophage  $\phi$ 6. *Journal of Virology* **73**, 2420–2424.
- Vennema, H., Poland, A., Foley, J. & Pedersen, N. C. (1998).** Feline infectious peritonitis viruses arise by mutation from endemic feline enteric coronaviruses. *Virology* **243**, 150–157.
- Wang, L., Junker, D., Hock, L., Ebiary, E. & Collisson, E. W. (1994).** Evolutionary implications of genetic variations in the S1 gene of infectious bronchitis virus. *Virus Research* **34**, 327–338.
- Weaver, S. C., Kang, W. L., Shirako, Y., Rumenapf, T., Strauss, E. G. & Strauss, J. H. (1997).** Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses. *Journal of Virology* **71**, 613–623.
- Weiller, G. F. (1998).** Phylogenetic Profiles: a graphical method for detecting genetic recombinations in homologous sequences. *Molecular Biology and Evolution* **15**, 326–335.
- Weiss, B. G. & Schlesinger, S. (1991).** Recombination between Sindbis virus RNAs. *Journal of Virology* **65**, 4017–4025.
- White, K. A. & Morris, T. J. (1994).** Recombination between defective tombusvirus RNAs generates functional hybrid genomes. *Proceedings of the National Academy of Sciences, USA* **91**, 3642–3646.
- Worobey, M., Rambaut, A. & Holmes, E. C. (1999).** Widespread intra-serotype recombination in natural populations of dengue virus. *Proceedings of the National Academy of Sciences, USA* **96**, 7352–7357.