

RNA Virus Populations as Quasispecies

J. J. HOLLAND¹, J. C. DE LA TORRE², and D. A. STEINHAUER³

1	Introduction	1
1.1	Mutation Frequencies and Rates of Evolution of RNA Viruses Are Generally Very High	2
1.2	RNA Virus Clones are Quasispecies Populations	2
2	Base Misinsertion Frequencies of Polymerases in the Absence of Proofreading and Repair Mechanisms	3
2.1	Tautomeric Equilibrium Constants and Polymerase Misinsertion Error	3
2.2	Thermodynamic Constraints on Base Pair Discrimination	5
2.3	Polymerase Misinsertion Frequencies In Vitro	6
3	Editing and Mismatch Repair: Do They Occur for RNA Molecules or for RNA Viruses Genomes?	7
4	Do Most (or All) RNA Viruses Produce Quasispecies Populations?	8
5	Population Biology and Evolution of RNA Virus Quasispecies	10
6	Implications of RNA Virus Quasispecies Populations for Pathogenesis and Disease	13
7	What Can Be Predicted Regarding the Future Evolution of RNA Viruses?	15
8	Summary	16
	References	16

1 Introduction

This chapter discusses the high mutation frequencies and rapid evolution potential of RNA viruses. The concepts discussed are applicable to all "ordinary" RNA viruses (riboviruses), viroids and satellite RNAs; to retroviruses; and to viruses (such as the hepadnaviruses) with DNA genomes which replicate via RNA transcripts. Because DNA virus polymerases can have proofreading (KORNBERG 1974), their mutation frequencies can be much lower than those of RNA viruses. For example, the mutation rate of bacteriophage T4 approximates 10^{-8} per base pair per replication (DRAKE 1969). However, some DNA viruses may avoid high-fidelity replication mechanisms (DRAKE et al. 1969; HALL et al. 1984) to

¹ Department of Biology and Institute of Molecular Genetics, University of California at San Diego, La Jolla, California 92093-0116, USA

² Department of Neuropharmacology, Research Institute of Scripps Clinic, La Jolla, California 92037, USA

³ National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, England

gain the evolutionary advantages of high mutation frequencies (SMITH and INGLIS 1987).

1.1 Mutation Frequencies and Rates of Evolution of RNA Viruses Are Generally Very High

Recent evidence indicates that most RNA viruses form complex "quasispecies" populations which can evolve very rapidly, due to extremely high mutation frequencies per site in RNA virus genomes (ranging between 10^{-3} and 10^{-6} and usually of the order of 10^{-4} – 10^{-5}). This evidence has been reviewed extensively and referenced in the following reviews and recent articles (HOLLAND et al. 1982, 1989, 1990, 1991; DOMINGO et al. 1985; SMITH and INGLIS 1987; ZIMMERN 1988; DOMINGO and HOLLAND 1988; COFFIN 1986, 1990; TEMIN 1989; PATHAK and TEMIN 1990; DE LA TORRE et al. 1990; DOMINGO 1989; EIGEN and BIEBRICHER 1988; CHAO 1990; MYERHANS et al. 1989; GROENINK et al. 1991; DOUGHERTY and TEMIN 1988; KEW et al. 1990; KINNUNEN et al. 1990; SMITH and PALESE 1988; STEINHAEUER and HOLLAND 1987; STEINHAEUER et al. 1989a, b; BILSEL and NICHOL 1990; BALFE et al. 1990; VARTANIAN et al. 1991; SKEHEL and WILEY 1988; BOEGE et al. 1991; IMAZEKI et al. 1990; HEINZ et al. 1989; STRAUSS and STRAUSS 1988; GEBAUER et al. 1988; VILLAVERDE et al. 1991; HAHN et al. 1989a, b; BURNS and DESROSIERS 1991; WARD et al. 1988).

1.2 RNA Virus Clones are Quasispecies Populations

Obviously if mutation frequencies usually exceed 10^{-5} at most sites in RNA virus genomes that average about 10 kb in size, then even clones of such viruses will inevitably consist of a complex mixture of different but related genomes, all of which must compete during replication of the clone and its progeny. Eigen, Schuster and their colleagues (EIGEN and SCHUSTER 1979; EIGEN et al. 1981; EIGEN and BIEBRICHER 1988) introduced the term "quasispecies" to refer to the diverse, rapidly evolving and competing RNA populations postulated as earliest life forms on earth, and the principles they have elaborated for quasispecies evolution apply well to RNA virus populations today (see for example DOMINGO et al. 1985; DOMINGO and HOLLAND 1988; GEBAUER et al. 1988; HOLLAND et al. 1989, 1990, 1991; PATHAK and TEMIN 1990; COFFIN 1990; DE LA TORRE et al. 1990; KURATH and PALUKAITIS 1990; GROENINK et al. 1991; BURNS and DESROSIERS 1991). For the discussion below, the term "consensus sequence" refers to the (average) sequence actually observed when aggregate genomes of a clonal (or non-clonal) virus population are sequenced. "Master sequence" refers to the most fit genome sequence (or sequences) within a complex quasispecies population replicating in a defined environment. "Mutant spectrum" refers to all of those competing virus variants which differ from the master sequence(s) and/or from the single genome which generated a quasispecies clone. Quasispecies populations are not hypothetical abstractions. They exist in animals and humans infected by RNA viruses such as vesicular stomatitis virus (VSV), foot-

and-mouth disease virus, poliovirus, human immunodeficiency virus (HIV)-1, influenza viruses, etc. as reviewed in the following chapters of this volume. Quasispecies populations also exist in clones of RNA viruses (including retroviruses) as shown by extreme mutation frequencies in clones and by rapid selection of mutants with increased fitness in the presence of antibody, drugs, new host cell types, etc. (See the following chapters, and the references cited above.)

The concept that RNA viruses are highly mutable and adaptable is not new. Very high mutation frequencies for phenotypes such as temperature sensitivity, pathogenicity, or plaque size mutants were reported decades ago (see for example GRANOFF 1961, 1964; FIELDS and JOKLIK 1969; EGGERS and TAMM 1965; PRINGLE 1970; KUNKEL 1947; THEILER and SMITH 1939). Indeed, the great adaptability of RNA viruses was utilized to prepare attenuated vaccines even before the genetic and biochemical nature of viruses was understood. More than 100 years ago Pasteur and his colleagues prepared the "fixed" vaccine strain of rabies virus by repeated passage of wild-type "street" virus intracerebrally in rabbits until it was no longer virulent for dogs by the subcutaneous route. Subsequently, a number of successful human attenuated virus vaccines have been prepared by multiple adaptive passages of virulent virus in new host cell types *in vitro* (THEILER and SMITH 1937; SABIN and BOULGER 1973; ENDERS *et al.* 1960). Despite this general knowledge that RNA viruses are quite mutable and adaptable, their extreme mutation frequencies have been documented only recently with the advent of rapid techniques for genome sequencing and mapping. Although many virologists are now aware of the quasispecies nature of populations of RNA viruses, their many biological implications are probably not generally appreciated. A number of authors have reported a quite low mutation frequency for some RNA viruses or some genome sites or genome regions of certain RNA viruses, and some examples of these will be discussed below following the general review of polymerase misincorporation frequencies. If it eventually can be rigorously proved that certain sites (or segments) of some RNA virus have very low mutation frequencies, this would not negate the quasispecies nature of populations of that virus if most other nucleotide sites (and segments) undergo mutation frequencies between 10^{-4} and 10^{-5} . As discussed below, polymerase error of the order of 10^{-4} – 10^{-5} is generally to be expected in the absence of proofreading and repair functions.

2 Base Misinsertion Frequencies of Polymerases in the Absence of Proofreading and Repair Mechanisms

2.1 Tautomeric Equilibrium Constants and Polymerase Misinsertion Error

The relative contributions of various mechanisms to DNA and RNA polymerase mutation frequencies are not completely established, and they clearly vary with

different polymerases. Most investigations have been done using DNA polymerases. The most elemental basis for inevitable polymerase errors is quantum mechanical. The Watson–Crick base pairing rules apply to the bases in their most favored tautomeric form, but shifts to rare tautomeric forms make mutations unavoidable during replication of both DNA and RNA. WATSON and CRICK (1953), in their original papers on the structure of DNA, suggested that base substitution mutations could be caused by the presence of minor tautomeric forms—for example, a shift of the hydrogen atom from the 6 amino position of adenine to the N1 position causing a mismatched pairing with cytosine. Quantum mechanical, spectral, thermodynamic, and kinetic analyses of various tautomeric forms of the bases and their tautomeric equilibria in aqueous solution have been laborious (see for example DANIELS 1972; KENNER et al. 1955; WOLFENDEN 1969; KWIATKOWSKI and PULLMAN 1975; DREYFUS et al. 1976), and are subject to uncertainties regarding the state of (and the effects of) base solvation or desolvation within the active sites of various polymerases (CULLIS and WOLFENDEN 1981). However, estimates of rare tautomer equilibrium constants of the order of 10^{-4} – 10^{-5} suggest a lower limit of about the same order of magnitude for base substitution frequencies at most sites in DNA and RNA. The classical paper by TOPAL and FRESCO (1976) utilized estimates of rare tautomer equilibria together with modeling of steric fit within a double helix to determine those rare tautomers able to be accommodated within a double helix in alternative base-pairing alignments. On this basis, they calculated likely base-mispairing frequencies, and compared their results with data available at that time. They proposed that rare tautomer mispairing (particularly of pyrimidine rare tautomers) should dictate transition frequencies at a level approximating 10^{-4} – 10^{-5} per site. However, transversions were suggested to require anti \rightarrow syn isomerization together with minor tautomeric forms to allow mispairing of purines, and hence were estimated to be about an order of magnitude less frequent than transitions. Because their model agreed well with transition and transversion frequencies of the mutD5 mutator strain of *Escherichia coli* they predicted that the mutD5 mutation probably eliminated the proofreading function of the DNA polymerase (TOPAL and FRESCO 1976). ECHOLS et al. (1983) later confirmed that mutD does lack the 3' \rightarrow 5'-exonucleolytic editing function (which is on the ϵ subunit of the DNA polymerase III holoenzyme), so the observed mutD5 mutation frequency is the basic unedited polymerase misinsertion frequency. Kinetic analyses of *E. coli* DNA polymerase I base misinsertion transitions in vitro (FERSHT et al. 1982, 1983) also generally agreed with the 10^{-4} – 10^{-5} frequencies predicted from tautomerization constant by TOPAL and FRESCO (1976). Their kinetic analysis also suggested that the polymerase does not greatly affect the tautomerization constant of active site-bound deoxynucleotide triphosphate (dNTP, as opposed to bases in the template). Specific misinsertions by some eukaryotic DNA polymerases such as chick γ polymerase are so low ($< 10^{-5}$) that these polymerases might be affecting tautomeric equilibria of some bases (KUNKEL and ALEXANDER 1986). Desolvation within the active site could be a factor. Also, protein complexes or DNA binding proteins

might enhance polymerase fidelity, perhaps by stabilizing templates (KUNKEL et al. 1979). Tautomer equilibrium constants provide only one of the reasons why minimal polymerase error frequency (in the absence of proofreading and repair) will generally exceed 10^{-5} .

2.2 Thermodynamic Constraints on Base Pair Discrimination

The pioneering studies of Kornberg and colleagues (KORNBERG 1974) demonstrated that there is a single active site in DNA polymerase at which all four dNTPs must compete for incorporation. This likely precludes the best possibility for extreme fidelity in which each specific dNTP in solution is bound to one of four specific sites and its major tautomeric form "confirmed" by steric fit before insertion opposite the one correct template base, the major tautomer of which was similarly confirmed. Early mutagenesis studies with base analogs demonstrated that base pairing is critical for much polymerase insertional specificity. This has been highlighted by the recent work of PICCIRILLI et al. (1990). They designed and synthesized two new "bases" designated kappa and pi. Although chemically distinct from natural bases, these were designed to form hydrogen-bonded base pairs in a DNA double helix. Kappa, which base pairs with xanthosine was able to direct incorporation of xanthosine opposite itself into both DNA and RNA. This exhibited approximately normal DNA and RNA polymerase misincorporation frequencies. It is clear here that base-pairing steric fit, rather than accurate base (and major tautomer) identification by polymerases is the chief mechanism for insertional specificity. However, base-pairing discrimination in aqueous solution involves major thermodynamic problems for polymerase fidelity as pointed out by LOEB and KUNKEL (1982). Free energy differences ($\Delta\Delta G$) between matched and mismatched base pairs in aqueous solution are estimated at no better than about 1–2 kcal/mol. This small difference in $\Delta\Delta G$ values cannot provide adequate base pair discrimination in solution to maintain base misinsertion frequencies per site better than about 10^{-1} – 10^{-2} . Nevertheless, as will be discussed below, unedited in vitro polymerase misincorporation frequencies per site generally average about 10^{-4} – 10^{-5} , and range between 10^{-3} to less than 10^{-5} (PRESTON et al. 1988). In contrast, a primitive lead-catalyzed, nonenzymatic polymerization of polyguanylate directed by a poly C template did exhibit the expected error levels, approximating 10^{-2} per site (LOHRMANN and ORGEL 1980). PETRUSKA et al. (1986, 1988) suggested that the $\Delta\Delta G$ s between matched and mismatched nucleotides are magnified at polymerase active sites by exclusion of water. Their studies suggested that *Drosophila* polymerase γ in vitro can amplify $\Delta\Delta G$ values more than ten fold (to approximately those expected in vacuo) by reducing entropy differences and increasing enthalpy differences. In this way, polymerase misincorporation frequencies can be lowered to about 10^{-4} – 10^{-5} per site, which is the level generally observed in vitro. Thus, for basic thermodynamic and quantum mechanical reasons, unedited polymerase mutation frequencies cannot ordinarily

be expected to be much lower than about 10^{-4} – 10^{-5} per site. Many other factors can also contribute to elevate base misinsertion frequencies above absolute minimal levels. Prominent among these are polymerase effects. Some polymerases exhibit generally high fidelity and others are more error prone. The latter are clearly not genetically designed to achieve the maximum possible base insertion fidelity, and their misinsertions are often favored at certain template sites.

2.3 Polymerase Misinsertion Frequencies In Vitro

It is not possible in this short chapter to review the extensive literature on mutagenesis. The reader is referred to the excellent recent reviews edited by MOSES and SUMMERS (1988). However, several important points relevant to this volume are outlined below.

In the absence of a proofreading function (or when the 3'→5'-exonuclease editing function is suppressed) *in vitro* polymerase error frequencies vary from as high as approximately 10^{-3} or 10^{-4} to as low as about 6×10^{-6} (LOEB and KUNKEL 1982; PRESTON et al. 1988). The latter value was obtained only with the four-subunit complex of calf thymus polymerase α -primase whereas homopolymeric preparations of the same polymerase α exhibited a misincorporation frequency between 10^{-4} and 10^{-5} . Therefore the presence of an undetected 3'→5'-exonuclease editing function could not be ruled out. Editing nucleases are often absent from polymerases of eukaryotic cells, but in some cases exogenous 3'→5'-exonuclease proteins may associate with these polymerase molecules *in vivo*, so the absence of editing functions on polymerases *in vitro* cannot disprove a role *in vivo*. Not only do different polymerases exhibit varying error rates *in vitro*, but they also can exhibit bias toward certain types of mismatch and preferentially misincorporate at or near certain template sequences. They also differentially incorporate certain base analogs, so template sequences and steric considerations at or near the active site are clearly important (LOEB and KUNKEL 1982; PRESTON et al. 1988; KUNKEL and ALEXANDER 1986; see also the chapter by WILLIAMS and LOEB).

Recent studies by KUNKEL and his colleagues (KUNKEL 1986; BEBENEK and KUNKEL 1990) suggest that base substitutions may often result from transient misalignment of the primer (nascent) strand on the template, and that base substitutions may, in a similar manner, initiate frameshift errors. There are many other factors which influence mutation frequencies and specificities. These include NTP substrate concentrations, metal ion variations, thermal effects on polymerase, base damage by oxidation, UV light, etc. Regardless of the multiplicity of factors which can be involved in mutagenesis, base misinsertion is an inevitable process with frequencies varying from about 10^{-3} to less than 10^{-5} and averaging about 10^{-4} – 10^{-5} . Editing and/or mismatch repair are required to achieve significantly higher fidelities.

3 Editing and Mismatch Repair: Do They Occur for RNA Molecules or for RNA Virus Genomes?

Proofreading (KORNBERG 1974) and mismatch repair are important processes in achieving fidelity in bacterial and phage DNA replication, but related processes are not presently as well characterized for eukaryotic cells (RADMAN and WAGNER 1986; LAHUE et al. 1989; MOSES and SUMMERS 1988). In eukaryotic systems, editing functions may be less tightly coupled with DNA polymerase than for bacteria and phages (i.e., they may be part of accessory proteins or protein complexes) and may therefore be more elusive. This makes it difficult to rule out their presence and function wherever they may not exist. Proofreading can improve DNA fidelity by orders of magnitude, and mismatch repair has the potential to achieve extreme fidelity via repeated scanning and repairing of DNA.

Mismatch repair is not possible for any RNA (or DNA) virus genomes if they are replicated and packaged as single-stranded molecules. This includes the vast majority of RNA viruses. Nor has proofreading yet been rigorously documented for any RNA molecules, viral or nonviral. There is no evidence to suggest the existence of 3' → 5'-exonuclease editing functions acting during synthesis of any RNA molecules. However it is difficult to disprove that they might exist on accessory protein molecules or complexes. Nevertheless, editing entails significant energy costs (FERSHT et al. 1982) and the expenditure of considerable energy to achieve high-fidelity RNA molecules should generally be counterproductive. The maximum error frequency of protein synthesis has been estimated to be between 10^{-3} and 10^{-4} (LOTFIELD and VANDERJAGT 1972; EDELMANN and GALLANT 1977; ELLIS and GALLANT 1982). This is a rather high fidelity in the complex process of translation. Apparently it includes kinetic proofreading by transfer RNA (tRNA) synthetases in aminoacylation (YAMANE and HOPFIELD 1977; JAKUBOWSKI 1990). With error frequencies exceeding 10^{-4} for translation, messenger RNA (mRNA) transcription fidelities need not exceed the usual fidelity levels observed for DNA polymerase insertions. In fact, measurements of RNA polymerase error frequencies (ROSENBERGER and HILTON 1983; BLANK et al. 1986) suggest RNA base misinsertion frequencies between 10^{-4} and 10^{-5} . The most careful measurements for wild-type *E. coli* RNA polymerase approximate 10^{-5} (BLANK et al. 1986) which exceeds by an order of magnitude the fidelity of translation. A mutant of *E. coli* RNA polymerase exhibited a lower fidelity (BLANK et al. 1986), but if any editing function is involved in *E. coli* polymerase transcription, it does not approach the high fidelity that can be provided by DNA polymerases with 3' → 5'-exonuclease. For example, the editing function associated with the ϵ subunit of the *po*/III DNA polymerase of *E. coli* can improve fidelity by up to 10^4 -fold (ECHOLS et al. 1983). It is not possible to exclude some low-fidelity proofreading in RNA synthesis such as the kinetic proofreading postulated by HOPFIELD (1974) (e.g., selective pyrophosphorolysis of mismatched bases). Pyrophosphorolysis does occur at significant levels during RNA transcription (KASSAVETIS et al. 1986).

One might argue that a structural (ribosomal) RNA could require higher fidelity. However, we have recently examined *in vivo* base misincorporation frequencies in ribosomal RNAs of mouse, hamster, and human cells, and base substitution frequencies greater than 10^{-4} per site were observed (STEINHAUER and HOLLAND, unpublished data). We have also been examining the virion-associated replicase of VSV and have as yet observed no evidence for any proofreading (STEINHAUER, DOMINGO, and HOLLAND, unpublished data). The only report of a proofreading-like activity for RNA is that of ISHIHAMA et al. (1986) for influenza virus RNA polymerase. However, this "editing" involved only removal of excess G residues added to cannibalized capped host cell mRNA primers. This was not inhibited by incorporation of 1-thiotriphosphates into RNA. Such inhibition is a hallmark of exonucleases of the 3' → 5' DNA proofreading (LOEB and KUNKEL 1982). All evidence presently available suggests that base misinsertion frequencies of RNA polymerases are similar for DNA polymerases. Proofreading is apparently absent or inefficient for most (or all) RNA synthesis, including viral RNA replication and reverse transcription by retroviruses (see the chapters by DOMINGO et al., COFFIN, WILLIAMS and LOEB, and WAIN HOBSON).

4 Do Most (or All) RNA Viruses Produce Quasispecies Populations?

Not all riboviruses nor retroviruses have yet been examined carefully to determine mutation frequencies and population heterogeneity. The great majority of presently available evidence suggests high mutation frequencies at most sites examined, so that even clonal populations must have heterogeneous quasi-species distributions.

Despite generally high mutation frequencies in most RNA viruses, it is important to remember that polymerase frequencies are known to vary, that for each polymerase there can be template sites at which misincorporation frequencies are considerably higher or lower than average, and that certain polymerase can favor certain specific mispairs. For example, eukaryotic polymerase β produces T·G mispairs at several orders of magnitude higher frequencies than does polymerase γ (KUNKEL and ALEXANDER 1986). Because RNA virus replicases and reverse transcriptases also exhibit preferences and variabilities there can be confusion regarding overall viral mutation frequencies. The following points should be emphasized:

1. Documentation of a low mutation frequency at a particular genome site of an RNA virus does *not* contradict evidence for high frequencies overall. For example, DURBIN and STOLLAR (1986) and HAHN et al. (1989a, b) reported three different specific site reversion frequencies in the Sindbis virus genome exhibiting mutation frequencies of the order of 10^{-7} to less than 10^{-8} yet HAHN

- et al. (1989a, b and back references) also observed many other Sindbis virus mutations occurring at frequencies between 10^{-3} and 10^{-5} suggesting that the three low-frequency reversion sites are rather unique, perhaps due to the sequence context in which they are located. STEC et al. (1986) likewise reported spontaneous mutation frequencies for Sindbis virus resistance to neutralizing monoclonal antibodies to be of the order of $10^{-3.5}$ – 10^{-5} . With poliovirus, SEDIVY et al. (1987) reported reversion frequencies for an amber mutant to be of the order of 10^{-6} whereas DE LA TORRE et al. (1990) reported specific single site mutation frequencies for poliovirus of the order of 10^{-4} .
2. Low mutation frequencies in only a few genome sites would not greatly alter population heterogeneity. In fact, if half of the sites in an RNA virus genome could be shown to have very low mutation frequencies, high frequencies at the other sites would still generate quasispecies populations.
 3. Extremely low mutation frequencies at a large number of genome sites in any RNA virus (i.e., frequencies of the order of 10^{-7} – 10^{-8}) should arouse suspicion of, and a search for, a proofreading function. As mentioned above, such a proofreading function need not necessarily be a part of the virus replicase, but could be due to an accessory protein or protein complex.
 4. Low mutation frequencies requiring a specific mismatch at a particular virus genome site do not even dictate a low *overall* mutation frequency at *that* site. For example, a phenotypic reversion requiring a T·T mismatch transversion at a site during genome replication might be of low frequency, but the same site could have very high transition frequencies. Also some low frequency (e.g., T·T) mismatching during genome replication might be increased considerably by a more favored (e.g., A·A) mismatching during antigenome replication (or vice versa). Strand replication asymmetry will increase or decrease such compensatory effects.
 5. RNA virus mutation frequencies of the order of 10^{-8} or less at any site are extremely unlikely to be due to high polymerase insertion fidelity. Frequencies this low are evidence either for an editing function, or for less interesting causes (e.g., the mutation or reversion requires several base substitutions—one at the site being examined and a necessary compensatory base substitution elsewhere; or the mutation assay allows negative selective forces to skew the observed frequency).
 6. If editing functions do exist which can reduce RNA replicase or reverse transcriptase mutation frequencies to 10^{-8} per site or lower, these have profound biological implications. Such editing functions, if they exist, are deserving of considerable effort toward purification and characterization.
 7. The RNA virus mutation frequencies observed (between 10^{-3} and 10^{-5}) at most RNA genome sites are so extreme that they approach the error threshold beyond which higher frequencies will lead to irreversible information loss and genome “error catastrophe” (EIGEN and BIEBRICHER 1988). In fact, chemical mutagenesis was unable to increase defined single site mutation frequencies in VSV and poliovirus more than 1.1- to 2.8- fold at any level of mutagenesis or virus survival level (HOLLAND et al. 1990). Clearly, RNA virus replicases and

reverse transcriptases can evolve to reduce their fidelity (and enhance their biological adaptability), but only to a limited extent compatible with viability and fitness; VSV mutants described by PRINGLE et al. (1981) and retrovirus "provirus hypermutants" described by PATHAK and TEMIN (1990) are examples.

8. Defective or partially defective (nonmaturing) viruses and, to a limited extent, infectious RNA viruses can exhibit specific biased "hypermutations" (CATTANEO et al. 1988; O'HARA et al. 1984; also see the chapters by CATTANEO and BILLETER). Such hypermutations may occur by various forms of RNA editing (SIMPSON 1990; LAMB and DREYFUSS 1989) such as the cellular unwinding activity of BASS et al. (1989) or less frequently by repetitious replicase errors (PATHAK and TEMIN 1990; BILSEL and NICHOL 1990; VARTANIAN et al. 1991).
9. Whenever sites in RNA virus genomes appear to have low mutation frequencies of the order of 10^{-6} – 10^{-8} , chemical mutagenesis should be able to increase greatly the mutation frequencies at such sites. This should occur without massive lethality due to error catastrophe (EIGEN and BIEBRICHER 1988; HOLLAND et al. 1990).

5 Population Biology and Evolution of RNA Virus Quasispecies

Because the mutation frequencies of RNA viruses exceed by more than a millionfold those of their eukaryotic hosts, extremely rapid virus evolution is anticipated and frequently observed (see reviews referenced in Sect. 1). A well-known example is the rapid continuous evolution of HIV-1 in infected humans (see the chapters by COFFIN, WILLIAMS and LOOB, WAIN-HOBSON, and DOOLITTLE and FENG. HIV is, of course, not unique in this respect since similar rapid evolution occurs in animals or humans naturally infected with foot- and-mouth disease virus (DOMINGO et al.), influenza virus (GORMAN et al.), poliovirus (KINNUNEN et al.), measles virus, (CATTANEO and BILLETER) and other viruses. However, relative long-term stasis of virus genomes can be observed in nature and in laboratory experiments. Such relative stasis does not imply that populations of these RNA viruses are not quasispecies, nor that their mutation frequencies are lower, nor that they are incapable of rapid evolution under different circumstances. For example, the genomes of eastern equine encephalitis virus in North America have exhibited relative stasis for the last half century while South American strains have apparently been evolving more rapidly (see chapter by WEAVER et al.). Some plant viruses also exhibit population stability (RODRIGUEZ-CEREZO et al. 1989). Likewise, influenza A virus, despite its obvious capacity for rapid evolution, can exhibit relative stasis of some genes (GORMAN et al. 1990; see chapter by GORMAN et al.). Finally, human T-cell lymphotropic virus (HTLV) appears to be evolving relatively slowly (INA and GOJOBORI 1990). Clones

of VSV under laboratory conditions can exhibit relative population stasis or extremely rapid evolution depending on passage conditions. Rapid evolution is promoted by conditions which lead to loss of population equilibrium (i.e., loss of dominance by previously most fit master sequences, and rise to dominance of new master sequences). Obviously, repeated environmental changes readily promote disequilibrium whether these changes are external (such as sequential infection of new hosts or host cell types) or internal (such as sequential interference due to the generation of changing populations of defective interfering virus particles in persistently infected cells). This is reviewed in HOLLAND et al. 1982; DOMINGO et al. 1985; DOMINGO and HOLLAND 1988; STEINHAUER et al. 1989).

It might seem paradoxical that heterogeneous quasispecies populations of RNA viruses can exhibit slow evolution or periods of stasis despite extreme mutation frequencies and rapid replication. However this can be explained by selection for fit master sequences in rather constant environments (EIGEN and BIEBRICHER 1988). The classical population biology theories of WRIGHT (1977, 1982) provide a useful mathematical paradigm for visualizing evolution of rather small populations (demes) in "adaptive landscapes." Random genetic variation coupled with environmental selection leads to random drift in an adaptive landscape (schematized in Fig. 1). Quantitative polygenic phenotypic characteristics are plotted on the Y and Z axes. Each combination of genetic characters has a mean fitness in a given environment, and this is plotted on the X, or fitness axis to provide an adaptive landscape. Small isolated, related populations can be represented as points on this landscape, and they will tend to spend a large fraction of evolutionary time near peaks of high fitness rather than less adaptive or nonadaptive ridges or valleys. Genetic variation and selection would tend to move a small population up a peak. With continuous selection, a population might become isolated on a peak even if there were much more highly adaptive peaks nearby representing better combinations of characters. However, WRIGHT (1977, 1982) envisioned landscapes with numerous peaks connected by

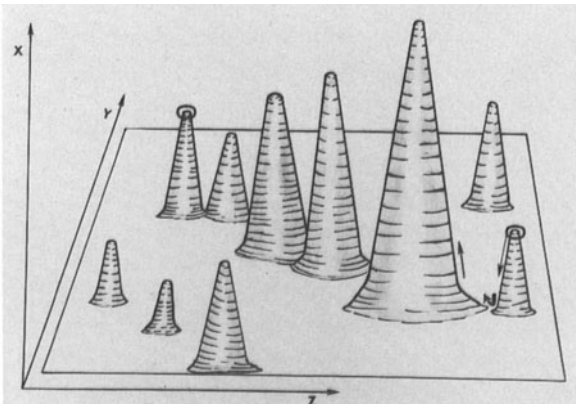


Fig. 1. Hypothetical adaptive landscape for RNA virus evolution in a constant adaptive environment. See text and WRIGHT (1977, 1982)

adaptive valleys so that random genetic drift could allow populations to move to other peaks. Once a population ascended to a very high fitness peak, gene flow would spread adaptive genes to other populations. WRIGHT (1982) suggested that his theories were relevant to recent paleobiological evidence for punctuated equilibrium (GOULD and ELDRIDGE 1977).

LANDE (1985) and NEWMAN et al. (1985) elegantly elaborated the applicability of WRIGHTS theories to punctuated equilibria during evolution. They showed that in a fixed, unchanging landscape, the expected time for population transitions between peaks should be extremely short. In contrast, the time spent near locally optimal peaks is long and increases approximately exponentially with effective population size (LANDE 1985). The timescale during transitions between peaks is short even though random genetic variations are small, and despite initial movement against selection during the transitions from one peak to another (NEWMAN et al. 1985). Thus, punctuated equilibria are explained even in a rather constant environment. Whenever the adaptive landscape changes abruptly as a result of a major environmental change, new selective forces will punctuate the equilibrium.

These theories are applicable to RNA virus populations even though virus populations can be extremely large, and virus mutation rates and evolution rates extremely high. In fact, these characteristics of RNA viruses should make them very useful for studies of population biology. Consider a quasispecies population of an RNA virus in the adaptive landscape of Fig. 1 to have ascended the lower right adaptive peak. The population may remain in equilibrium around the top of this peak even though there is a much higher adaptive optimum nearby. This will occur particularly when the local virus population remains large. Very high mutation frequencies can counter this by accumulating a small proportion of variants having numerous mutations so that there can be movement down the peak (against selection), but the more fit variants near the top (close to the master sequences [s]) will dominate. Thus, population equilibrium might be maintained for relatively long periods of time despite high mutation frequencies. In fact, DE LA TORRE et al. (1990) observed that a mutant of VSV of vastly superior fitness could not rise to dominate its diverse quasispecies progenitor population of lower mean fitness unless it was seeded above a critical threshold level, and unless at least some intracellular replication occurred in the absence of competitor variants (by carrying out dilute passages during competition).

Thus, in a constant environment, a quasispecies variant swarm might hover for a relatively long period near the moderately adaptive peak depicted at the lower right hand of Fig. 1. Eventually, when a low probability accumulation of many appropriate mutations moves a subset of the quasispecies population down the peak (against selection) and across a nonadaptive valley or ridge (arrows in Fig. 1), strong positive selection should quickly move the population up the adjacent highly adaptive peak to produce new, highly fit master (and consensus) sequences together with a mutant spectrum of higher average fitness. This movement to the new peak must occur *rapidly*, and *only* when virus transmission leads to a rather low population (LANDE 1985; NEWMAN et al. 1985).

Virus transmission from host to host, or from one area to another within a host, often involves small virus populations or even a single virus particle (genetic bottleneck transmission).

Finally, the fastest way to disrupt a stable virus equilibrium near a highly adaptive peak is to change the adaptive landscape. This happens frequently with viruses (e.g., during immune responses; changes of host species or of cell type within a single host; interference by defective viruses, inflammatory responses, etc.). Hence, rapid evolution of RNA viruses is often more evident than is relative evolutionary stasis, but both do occur (see the chapters by DOMINGO et al., GORMAN et al., WEAVER et al., COFFIN, and DOOLITTLE and FENG). WRIGHTS two-dimensional combinations of characters are of course an oversimplification for viruses which, intracellularly, compete (and interact) with the countless mutants (and mutant gene products) which they regularly generate. Still this paradigm can give useful insights into the complexities of RNA virus evolution. RNA viruses should provide good laboratory models for evolution and evolutionary theories.

6 Implications of RNA Virus Quasispecies Populations for Pathogenesis and Disease

Many of the consequences of RNA virus quasispecies virus populations are rather obvious. For example, RNA genome population diversity can allow rapid evolution, antigenic change and immune system evasion, and rapid development of effective resistance to antiviral drugs (see the chapters by ; DOMINGO 1990). Such diversity also facilitates and complicates the establishment and progression of chronic infections such as acquired immunodeficiency syndrome (AIDS) or SSPE. The presence of a dynamic quasispecies population of RNA genomes in infected individuals can make it difficult to explore fully the pathogenesis of infections. This is fairly obvious in the case of AIDS syndromes (see below), but it is true to some extent for all RNA virus infections. A heterogeneous swarm of RNA genomes in an infected tissue or organ renders it impossible to assign specific pathological significance to any consensus sequence recovered from the diseased tissue, and, of course, the master sequences and consensus sequences continuously evolve during disease progression and/or recovery. For example, in mice infected with the Armstrong strain of lymphocytic choriomeningitis virus (LCMV), variants arise at high frequency which cause selective immunosuppression of those T cells designed to recognize and reject LCMV (AHMED et al. 1984). BOZZOW and OLDSTONE (personal communication) also observed a high frequency of LCMV variants with differing interactions with cellular receptors. Finally, OLDSTONE (personal communication) recently found a high frequency of clones capable of causing the mouse growth hormone "stunting" syndrome within the wild-type quasi-

species population of the WE strain of LCMV. Since this wild-type WE strain of LCMV does not cause the growth hormone syndrome in persistently infected mice, it is clear that the complex quasispecies population of WE strain LCMV suppresses and Masks an important disease potential of genomes buried within its mutant spectrum. Quasispecies populations inevitably dictate a certain degree of uncertainty regarding the biology of all RNA virus diseases. In most cases, this should pose no practical problems for diagnosis, treatment, vaccination, etc., but in many cases it will. There is an unspoken assumption among many physicians and scientists that *a particular RNA virus will generally cause a particular disease*. This assumption may be true in a very broad practical sense, but it is important to understand that it can never be true in a formal scientific sense. Because *a particular RNA virus* simply does not exist, *a particular RNA virus disease* does not exist either. The science of infectious diseases is still in its infancy because we still understand so little of the fine details of host-pathogen interactions.

Any quasispecies swarm of RNA virus genomes in an infected human creates a quasispecies swarm of virus-encoded proteins, both intracellularly and extracellularly. As was discussed previously, all human (and other) protein pools must contain amino acid misincorporations at a level greater than 10^{-4} , so these could be considered to have a quasispecies type distribution. However, this should not be confused with the *replicating* quasispecies mutant swarms of RNA viruses. The latter selectively replicate and *amplify* certain subsets of their (nonmaster sequence) mutant spectrum in a completely unpredictable manner, thereby creating significant levels of many nonmaster, nonconsensus sequence proteins. During an influenza A virus, or measles virus or rhinovirus outbreak, for example, no two infected humans will encounter the same array of foreign virus-encoded proteins. Therefore, the acute effects and subtle chronic effects of infections will differ not only because we all vary genetically, physiologically and immunologically, but also because we all experience a different array of quasispecies challenges. These facts are easily overlooked by clinicians and scientists because disease syndromes are often grossly similar for each type of virus, and because it would appear to make no difference in a practical sense. However, for the person who develops Guillain–Barre syndrome following a common cold, or for the individual who remains healthy despite many years of HIV-1 infection, for example, it may make all the difference in the world.

The extent to which chronic degenerative diseases and autoimmune diseases of man are triggered by exogenous pathogens is still unknown. It is clear that virus proteins can exhibit molecular mimicry (FUJINAMI and OLDSTONE 1989) and can induce cell-mediated autoimmunity (TER MEULEN 1989; KYUWA et al. 1991). In the case of RNA virus infections, there is no reason to expect that the ability to trigger degenerative and autoimmune diseases of man will be confined solely to those proteins or protein domains encoded by master sequence or consensus sequence genomes. The former may often represent a small percentage of total virus encoded proteins and the latter often be vanishingly small. The consensus sequence is an average of the most abundant nucleotide present at each

position of all genomes in the quasispecies swarm and it need not necessarily even exist in any one genome present.

Similar arguments apply to viral cell tropisms, and to cell and tissue pathology in infected humans. It can be extremely difficult or impossible to ascribe subtle cell tropism differences or cellular pathology to the consensus sequence or to any "clonal" sequences recovered from a patient's cells and tissues. It has recently been observed that even the relative fitness of RNA virus clones can change rapidly, and that even clonal populations adapt rapidly to new host cell types (HOLLAND et al. 1991; MARTINEZ et al. 1991). It is important (if disturbing) to remember that clonal sequences of viruses recovered from patients can be distorted by the methods employed for recovery and amplification, whether these are growth of virus plaques, reverse transcription followed by polymerase chain reaction (PCR), or whatever. Thus RNA virology and a certain level of uncertainty go hand in hand. Finally, the recent identification of superantigens encoded by proviruses of retroviruses (MARRACK et al. 1991; DYSON et al. 1991; FRANKEL et al. 1991; WOODLAND et al. 1991) extends these concepts to T cell stimulation or deletion and immunosuppression. Any replicating RNA virus, whether retrovirus or ribovirus, might sometimes produce protein sequences with superantigen properties. When these are encoded by selected, stable, integrated proviruses they can be clearly identified (as has been done) but when such superantigens are hidden within quasispecies swarms, their presence, and their role in disease, may generally go undetected.

7 What Can Be Predicted Regarding the Future Evolution of RNA Viruses?

Very little can be predicted, except that rapid virus evolution will continue. Some "new" human RNA virus diseases will inevitably emerge, and some older diseases will disappear or become less significant. Although some RNA viruses can achieve a rather stable equilibrium in certain environments, most will generally evolve at rates exceeding by a millionfold those of their eukaryotic hosts. Because mutation is probabilistic and produces immense and largely indeterminate quasispecies RNA virus population, the directions of RNA genome evolution must remain unpredictable.

Although new RNA virus diseases of humans will continue to emerge at indeterminate intervals, the viruses themselves will not really be new, but rather mutated and rearranged to allow infection of new hosts, or to cause new disease patterns. It is important to remember that every quasispecies genome swarm in an infected individual is unique and "new" in the sense that no identical population of RNA genomes has ever existed before and none such will ever exist again. Because a large fraction of sites on RNA virus genomes can eventually be substituted and because genomes can also be expanded, contracted, recom-

bined and rearranged, the number of potential genome permutations which must be tried to find all viable alternatives is essentially infinite in the space time of our universe. As human populations continue to grow exponentially, the number of ecological niches for human RNA virus evolution grows apace and "new" human virus outbreaks will likely increase apace. Most "new" human viruses will be unremarkable—that is they will generally resemble "old" ones. Inevitably, some will be quite remarkable, and quite undesirable. When discussing RNA virus evolution, to call an outbreak (such as AIDS) "remarkable" is merely to state that it is of lower probability than an "unremarkable" outbreak.

8 Summary

RNA virus mutation frequencies generally approach maximum tolerable levels, and create complex indeterminate quasispecies populations in infected hosts. This usually favors extreme rates of evolution, although periods of relative stasis or equilibrium, punctuated by rapid change may also occur (as for other life forms). Because complex quasispecies populations of RNA viruses arise probabilistically and differentially in every host, their compositions and exact roles in disease pathogenesis are indeterminate and their directions of evolution, and the nature and timing of "new" virus outbreaks are unpredictable.

References

- Ahmed R, Salmi A, Butler LD, Chiller JM, Oldstone MBA (1984) Selection of genetic variants of LCMV in spleens of persistently infected mice: role in suppression of cytotoxic T lymphocyte response and viral persistence. *J Exp Med* 160: 521–540
- Balfe P, Simmonds P, Ludlam CA, Bishop JO, Leigh Brown AJ (1990) Concurrent evolution of human immunodeficiency virus type 1 in patients infected from the same source: rate of sequence change and low frequency of inactivating mutations. *J Virol* 64: 6221–6233
- Bass BL, Weintraub H, Cattaneo R, Billeter MA (1989) Biased hypermutation of viral genomes could be due to unwinding/modification of double stranded RNA. *Cell* 56: 331
- Batschelet E, Domingo E, Weissman C (1976) The proportion of revertant and mutant phage in a growing population as a function of mutation and growth rate. *Gene* 1: 27–32
- Bebenek K, Kunkel TA (1990) Frameshift errors initiated by nucleotide misincorporation. *Proc Natl Acad Sci USA* 87: 4946–4950
- Binsel PA, Nichol ST (1990) Polymerase errors accumulating during natural evolution of the glycoprotein gene of vesicular stomatitis virus Indiana serotype isolates. *J Virol* 64: 4873–4883
- Blank A, Gallant JA, Burgess RR, Loeb LA (1986) A RNA polymerase mutant with reduced accuracy of chain elongation. *Biochemistry* 25: 5920–5928
- Boege U, Kobasa D, Onodera S, Parks GD, Palmenberg AC, Scraba DG (1991) Characterization of mengo virus neutralization epitopes. *Virology* 181: 1–13
- Burns DPW, Desrosiers RC (1991) Selection of genetic variants of simian immunodeficiency virus in persistently infected rhesus monkeys. *J Virol* 65: 1843–1854
- Cattaneo R, Schmid A, Eschle D, Baczko K, ter Meulen V, Billeter MA (1988) Biased hypermutation and other genetic changes in defective measles viruses in human brain infections. *Cell* 55: 255–265

- Chao L (1990) Fitness of RNA virus decreased by Muller's ratchet. *Nature* 348: 454–455
- Coffin JM (1986) Genetic variation in AIDS viruses. *Cell* 46: 1–4
- Coffin JM (1990) Genetic variation in retroviruses. In: Kurstak E, Marusyk RG, Murphy FA, van Regenmortel MHV (eds) *Applied virology research*, vol 2. Plenum, New York, pp 11–33
- Cullis PM, Wolfenden R (1981) Affinities of nucleic acid bases for solvent water. *Biochemistry* 20: 3024–3028
- Daniels M (1972) Tautomerism of uracil and thymine in aqueous solution; spectroscopic evidence. *Proc Natl Acad Sci USA* 69: 2488–2491
- De La Torre JC, Wimmer E, Holland JJ (1990) Very high frequency of reversion to guanidine resistance in clonal pools of guanidine-dependent type I poliovirus. *J Virol* 64: 664–671
- Domingo E (1989) RNA virus evolution and the control of viral disease. *Prog Drug Res* 33: 93–133
- Domingo E, Holland JJ (1988) High error rates, population equilibrium and evolution of RNA replication systems. In: Domingo E, Holland JJ, Ahlquist P (eds) *RNA genetics*, vol 3. CRC, Boca Raton, pp 3–36
- Domingo E, Martinez-Salas E, Sobrino F, De la Torre JC, Portela A, Ortin J, Lopez-Galindez C, Pérez-Brena P, Villanueva N, Nájera R, VandePol S, Steinhauer D, DePolo N, Holland JJ (1985) The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance—a review. *Gene* 40: 1–8
- Dougherty JP, Temin HM (1988) Determination of the rate of base-pair substitution and insertion mutations in retrovirus replication. *J Virol* 62: 2817–2822
- Drake JW (1969) Comparative rates of spontaneous mutation. *Nature* 221: 1132
- Drake JW, Allen EF, Forsberg SA, Preparata RM, Greening EO (1969) Spontaneous mutation. *Nature* 221: 1128–1132
- Dreyfus M, Bensaude O, Dodin G, Dubois JE (1975) Tautomerism in cytosine and 3-methylcytosine. A thermodynamic and kinetic study. *J Am Chem Soc* 98: 6338–6349
- Durbin RK, Stollar V (1986) Sequence analysis of the E1 gene of a hyperglycosylated, host restricted mutant of Sindbis virus and estimation of mutation rate from frequency of revertants. *Virology* 154: 135–143
- Dyson PJ, Knight AM, Fairchild S, Simpson E, Tomonari K (1991) Genes encoding ligands for deletion of V β 11 T cells cosegregate with mammary tumour virus genomes. *Nature* 349: 531–532
- Echols H, Lu C, Burgers PMJ (1983) Mutator strains of *Escherichia coli*, *mutD* and *dnaQ*, with defective exonucleolytic editing by DNA polymerase III holoenzyme. *Proc Natl Acad Sci USA* 80: 2189–2192
- Edelman P, Gallant J (1977) Mistranslation in *E. coli*. *Cell* 10: 131–137
- Eggers HJ, Tamm I (1965) Coxsackie A9 virus: mutation from drug dependence to drug resistance. *Science* 148: 97–98
- Eigen M, Biebricher CK (1988) Sequence space and quasispecies distribution. In: Domingo E, Holland JJ, Ahlquist P (eds) *RNA genetics*, vol 3. CRC, Boca Raton, pp 211–245
- Eigen M, Schuster P (1979) *The hypercycle. A principle of natural self-organization*. Springer, Berlin Heidelberg New York
- Eigen M, Gardiner W, Schuster P, Winkler-Oswatitsch R (1981) The origin of genetic information. *Sci Am* 244: 88–188
- Ellis N, Gallant J (1982) An estimate of the global error frequency in translation. *Mol Gen Genet* 188: 169–172
- Enders J, Katz SL, Milovanovic MJ, Holloway A (1960) Studies on an attenuated measles virus vaccine. I. Development and preparation of vaccine: techniques for assay of effects of vaccination. *N Engl J Med* 263: 153–159
- Fersht AR, Knill-Jones JW, Tsui W-C (1982) Kinetic basis of spontaneous mutation. *J Mol Biol* 156: 37–51
- Fersht AR, Shi J-P, Tsui W-C (1983) Kinetics of base misinsertion by DNA polymerase I of *Escherichia coli*. *J Mol Biol* 165: 655–667
- Fields BN, Joklik WK (1969) Isolation and preliminary characterization of temperature sensitive mutants of retrovirus. *Virology* 37: 335–342
- Frankel WN, Rudy C, Coffin JM, Huber BT (1991) Linkage of *M/s* genes to endogenous mammary tumour viruses of inbred mice. *Nature* 349: 526–528
- Fujinami RS, Oldstone MBA (1989) Molecular mimicry as a mechanism for virus-induced autoimmunity. *Immunol Res* 8: 3–15
- Gebauer F, De La Torre JC, Gomes I, Mateu MG, Barahona H, Tiraboschi B, Bergmann I, Auge De Mello P, Domingo E (1988) Rapid selection of genetic and antigenic variants of foot- and-mouth disease virus during persistence in cattle. *J Virol* 62: 2041–2049

- Gorman OT, Bean WJ, Kawaoka Y, Webster RG (1990) Evolution of the nucleoprotein gene of influenza A virus. *J Virol* 64: 1487–1497
- Gould SJ, Eldridge N (1977) Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleontology* 3: 115–151
- Granoff A (1961) Induction of Newcastle disease virus mutants with nitrous acid. *Virology* 13: 402–408
- Granoff A (1964) Nature of the Newcastle disease virus population. In: Hanson RP (ed) Newcastle disease virus, an evolving pathogen. University of Wisconsin Press, Madison, pp 107–108
- Groenink M, Fouchier RAM, de Goede REY, de Wolf F, Gruters RA, Cuypers HTM, Huisman HG, Tersmette M (1991) Phenotypic heterogeneity in a panel of infectious molecular human immunodeficiency virus type 1 clones derived from a single individual. *J Virol* 65: 1968–1975
- Hahn YS, Strauss EG, Strauss JH (1989a) Mapping of RNA⁻ temperature sensitive mutants of sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *J Virol* 63: 3142–3150
- Hahn CS, Rice CM, Strauss EG, Lenches EM, Strauss JH (1989b) Sindbis virus ts103 has a mutation in glycoprotein E2 that leads to defective assembly of virions. *J Virol* 63: 3459–3465
- Hall JD, Coen DM, Fisher BL, Weisslitz M, Randall S, Almy RE, Gelep TP, Schaffer PA (1984) Generation of genetic diversity in herpes simplex virus: an antimutator phenotype maps to the DNA polymerase locus. *Virology* 132: 26–37
- Heinz BA, Rueckert RR, Shepard DA, Dutko FJ, McKinlay MA, Fancher M, Rossmann MG, Badger J, Smith TJ (1989) Genetic and molecular analyses of spontaneous mutants of human rhinovirus 14 that are resistant to an antiviral compounds. *J Virol* 63: 2476–2485
- Holland JJ, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S (1982) Rapid evolution of RNA genomes. *Science* 215: 1577–1585
- Holland JJ, De La Torre JC, Steinhauer DA, Clarke D, Duarte E, Domingo E (1989) Virus mutation frequencies can be greatly underestimated by monoclonal antibody neutralization of virions. *J Virol* 63: 5030–5036
- Holland JJ, Domingo E, De La Torre JC, Steinhauer DA (1990) Mutation frequencies at defined single codon sites in vesicular stomatitis virus can be increased only slightly by chemical mutagenesis. *J Virol* 64: 3960–3962
- Holland JJ, De La Torre JC, Clarke DK, Duarte E (1991) Quantitation of relative fitness and great adaptability of clonal populations of RNA viruses. *J Virol* (in press)
- Hopfield JJ (1974) Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci USA* 71: 4135–4139
- Imazeki F, Omata M, Ohto M (1990) Heterogeneity and evolution rates of delta virus RNA sequences. *J Virol* 64: 5594–5599
- Ina Y, Gojobori T (1990) Molecular evolution of human T-cell leukemia virus. *J Mol Evol* 31: 493–499
- Ishihama A, Mizumoto K, Kawakami K, Kato A, Honda A (1986) Proofreading function associated with the RNA-dependent RNA polymerase from influenza virus. *J Biol Chem* 261: 10417–10421
- Jakubowski H (1990) Proofreading in vivo: editing of homocysteine by methionyl-tRNA synthetase in *Escherichia coli*. *Proc Natl Acad Sci USA* 87: 4504–4508
- Kassavetis GA, Zentner PG, Geiduschek EP (1986) Transcription at bacteriophage T4 variant late promoters: an application of a newly devised promoter-mapping method. *J Biol Chem* 261: 14256–14265
- Kenner GW, Reese CB, Sir Todd AR (1955) The acylation of 3-methylcytosine. *J Chem Soc* 50C: 855–859
- Kwe OM, Nottay BK, Rico-Hess R, Pollansch MA (1990) Molecular epidemiology of wild poliovirus transmission. In: Kurstak E (ed) Applied virology research, vol 2. Plenum, New York, pp 199–220
- Kinnunen L, Huovilainen A, Pöyry T, Hovi T (1990) Rapid molecular evolution of wild type 3 poliovirus during infection in individual hosts. *J Gen Virol* 71: 317–324
- Kornberg A (1974) DNA synthesis. Freeman, San Francisco, pp 67–121
- Kunkel LO (1947) Variation in phytopathogenic viruses. *Annu Rev Microbiol* 1: 85–100
- Kunkel TA (1986) Frameshift mutagenesis by eucaryotic DNA polymerases in vitro. *J Biol Chem* 261: 13581–13587
- Kunkel TA, Alexander PS (1986) The base substitution fidelity of eucaryotic DNA polymerases. *J Biol Chem* 261: 160–166
- Kunkel TA, Meyer RR, Loeb LA (1979) Single-strand binding protein enhances fidelity of DNA synthesis in vitro. *Proc Natl Acad Sci USA* 76: 6331–6335
- Kurath G, Palukaitis P (1990) Serial passage of infectious transcripts of a cucumber mosaic virus satellite RNA clone results in sequence heterogeneity. *Virology* 176: 8–15

- Kwiatkowski JS, Pullman B (1975) Tautomerism and electronic structure of biological pyrimidines. *Adv Heterocycl Chem* 18: 199–335
- Kyuwa S, Yamaguchi K, Toyoda Y, Fujiwara K (1991) Induction of self-reactive T cells after murine coronavirus infection. *J Virol* 65: 1789–1795
- Lahue RS, Au KG, Modrich P (1989) DNA mismatch correction in a defined system. *Science* 245: 160–164
- Lamb RA, Dreyfuss G (1989) RNA structure. Unwinding with a vengeance [news]. *Nature* 337: 19–20
- Lande R (1985) Expected time for random genetic drift of a population between stable phenotypic states. *Proc Natl Acad Sci USA* 82: 7641–7645
- Loeb LA, Kunkel TA (1982) Fidelity of DNA synthesis. *Annu Rev Biochem* 52: 429–457
- Lotfield RB, Vanderjagt D (1972) The frequency of errors in protein biosynthesis. *Biochem J* 128: 1353–1356
- Lohrmann R, Orgel LE (1980) Efficient catalysis of polycytidylic acid-directed oligoguanylate formation by Pb^{2+} . *J Mol Biol* 142: 555–567
- Marrack P, Kushnir E, Kappler J (1991) A maternally inherited superantigen encoded by a mammary tumour virus. *Nature* 349: 524–526
- Martinez MA, Carillo C, Gonzalez-Candelaz F, Moya A, Domingo E, Sobrino F (1991) Fitness alteration of foot- and-mouth disease virus mutants; measurement of adaptability of viral quasispecies. *J Virol* 65: 3954–3957
- Moses RE, Summers WC (eds) (1988) DNA replication and mutagenesis. American Society for Microbiology, Washington, p 515
- Myerhans A, Cheynier R, Albert J, Seth M, Kwok S, Sninsky J, Morfeldt-Manson L, Åsjö B, Wain-Hobson S (1989) Temporal fluctuations in HIV quasispecies *in vivo* are not reflected by sequential HIV isolations. *Cell* 58: 901–910
- Newman CM, Cohen JE, Kipnis C (1985) New-darwinian evolution implies punctuated equilibria. *Nature* 315: 400–401
- O'Hara PJ, Nichol ST, Horodyski FM, Holland JJ (1984) Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell* 36: 915–924
- Pathak VK, Temin HM (1990) Broad spectrum of *in vivo* forward mutations, hypermutations, and mutational hotspots in a retroviral shuttle vector after a single replication cycle: substitutions, frameshifts, and hypermutations. *Proc Natl Acad Sci USA* 87: 6019–6023
- Petruska J, Sowers LC, Goodman MF (1986) Comparison of nucleotide interactions in water, proteins, and vacuum: model for DNA polymerase fidelity. *Proc Natl Acad Sci USA* 83: 1559–1562
- Petruska J, Goodman MF, Boosalis MS, Sowers LC, Cheong C, Tinoco I Jr (1988) Comparison between DNA melting thermodynamics and DNA polymerase fidelity. *Proc Natl Acad Sci USA* 85: 6252–6256
- Piccirilli JA, Krauch T, Moroney SE, Benner SA (1990) Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet. *Nature* 343: 33–37
- Preston BD, Zakour RA, Singer B, Loeb LA (1988) Fidelity of base selection by DNA polymerases. Site-specific incorporation of base analogs. In: Moses RE, Summers WC (eds) DNA replication and mutagenesis. American Society for Microbiology, Washington, pp 196–207
- Pringle CR (1970) Genetic characteristics of conditional lethal mutants of vesicular stomatitis virus induced by 5-fluorouracil, 5-azacytidine, and ethyl methanesulfonate. *J Virol* 5: 559–567
- Pringle CR, Devine V, Wilkie M, Preston CM, Dohn A, McGeoch DJ (1981) Enhanced mutability associated with a temperature-sensitive mutant of vesicular stomatitis virus. *J Virol* 39: 377–389
- Radman M, Wagner R (1986) Mismatch repair in *Escherichia coli*. *Annu Rev Genet* 20: 523–538
- Rodríguez-Cerezo E, Moya A, García-Arenal F (1989) Variability and evolution of the plant RNA virus pepper mild mottle virus. *J Virol* 63: 2198–2203
- Rosenberger RF, Hilton J (1983) The frequency of transcriptional and translational errors at nonsense codons in the *lacZ* gene of *Escherichia coli*. *Mol Gen Genet* 191: 207–212
- Sabin AB, Boulger LR (1973) History of Sabin attenuated poliovirus oral live vaccine strains. *J Biol Stand* 1: 115–118
- Sedivy JM, Capone JP, Raj Bhandary UL, Sharp PA (1987) An inducible mammalian amber suppressor: propagation of a poliovirus mutant. *Cell* 50: 379–389
- Simpson L (1990) RNA editing—a novel genetic phenomenon? *Science* 250: 512–513
- Skehel JJ, Wiley DC (1988) Antigenic variation in influenza virus hemagglutinin. In: Domingo E, Holland JJ, Ahlquist P (eds) RNA genetics, vol 3. CRC, Boca Raton, pp 139–146
- Smith BD, Inglis SC (1987) The mutation rate and variability of eukaryotic viruses: an analytical review. *J Gen Virol* 68: 2729–2740

- Smith FI, Palese P (1988) Influenza virus: high rates of mutation and evolution. In: Domingo E, Holland JJ, Ahlquist P (eds) RNA genetics, vol 3. CRC, Boca Raton, pp 123–135
- Stec DS, Waddell A, Schmaljohn CS, Cole GA, Schmaljohn AL (1986) Antibody-selected variation and reversion in Sindbis virus neutralization epitopes. *J Virol* 57: 715–720
- Steinhauer DA, Holland JJ (1987) Rapid evolution of RNA viruses. *Annu Rev Microbiol* 41: 409–433
- Steinhauer DA, de la Torre JC, Holland JJ (1989a) High nucleotide substitution error frequencies in clonal pools of vesicular stomatitis virus. *J Virol* 63: 2063–2071
- Steinhauer DA, de la Torre JC, Meier E, Holland JJ (1989b) Extreme heterogeneity in populations of vesicular stomatitis virus. *J Virol* 63: 2072–2080
- Strauss JH, Strauss EG (1988) Evolution of RNA viruses. *Annu Rev Microbiol* 42: 657–683
- Temin HM (1989) Is HIV unique or merely different? *AIDS* 2: 1–9
- Ter Meulen V (1989) Virus-induced, cell mediated immunity. In: Notkins AL, Oldstone MBA (eds) Concepts in viral pathogenesis, vol 3. Springer, Berlin Heidelberg, New York, pp 297–310
- Theiler M, Smith HH (1937) Use of yellow fever virus modified by in vitro cultivation for human immunization. *J Exp Med* 65: 787–800
- Theiler M, Smith HH (1939) The effect of prolonged cultivation in vitro upon the pathogenicity of yellow fever virus. *J Exp Med* 65: 767–787
- Topal MD, Fresco JR (1976) Complementary base pairing and the origin of substitution mutations. *Nature* 263: 285–289
- Vartanian J-P, Meyerhans A, Åsjö B, Wain-Hobson S (1991) Selection, recombination, and G→A hypermutation of human immunodeficiency virus type 1 genomes. *J Virol* 65: 1779–1788
- Villaverde A, Martinez MA, Sobrino F, Dopazo J, Moya A, Domingo E (1991) Fixation of mutations at the VP1 gene of foot- and-mouth disease virus. Can quasispecies define a transient molecular clock? *Gene* (in press)
- Ward CD, Stokes MAM, Flanagan JB (1988) Direct measurement of the poliovirus RNA polymerase error frequency in vitro. *J Virol* 62: 558–562
- Watson JD, Crick FHC (1953) The structure of DNA. *Cold Spring Harbour Symp Quant Biol* 48: 123–133
- Wolfenden RV (1969) Tautomerism equilibria in inosine and adenosine. *J Mol Biol* 40: 307–310
- Woodland DL, Happ MP, Gollob KJ, Palmer E (1991) An endogenous retrovirus mediating deletion of $\alpha\beta$ T cells? *Nature* 349: 529–530
- Wright S (1977) Experimental results and evolutionary deductions. In: Wright S (ed) Evolution and the genetics of populations, vol 3. Chicago, Chicago, pp 443–473
- Wright S (1982) Character change, speciation, and the higher taxa. *Evolution* 36: 427–443
- Yamane T, Hopfield JJ (1977) Experimental evidence for kinetic proofreading in the aminoacylation of tRNA by synthetase. *Proc Natl Acad Sci USA* 74: 2246–2250
- Zimmern D (1988) Evolution of RNA viruses. In: Domingo E, Holland JJ, Ahlquist P (eds) RNA genetics, vol 3. CRC, Boca Raton, pp 211–240