

From Quasispecies Theory to Viral Quasispecies: How Complexity has Permeated Virology

E. Domingo^{1,2,3*}, C. Perales^{1,3}

¹ Centro de Biología Molecular “Severo Ochoa” (CSIC-UAM)
Campus de Cantoblanco 28049, Madrid, Spain

² Centro de Astrobiología (CSIC-INTA), Instituto Nacional de Técnica Aeroespacial
Ctra. de Torrejón a Ajalvir, Km 4, 28850 Torrejón de Ardoz, Madrid, Spain

³ Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas
(CIBERehd), 08036 Barcelona, Spain

Abstract. RNA viruses replicate as complex and dynamic mutant distributions. They are termed viral quasispecies, in recognition of the fundamental contribution of quasispecies theory in our understanding of error-prone replicative entities. Viral quasispecies have launched a fertile field of transdisciplinary research, both experimental and theoretical. Here we review the origin and some implications of the quasispecies concept, with emphasis on internal interactions among components of the same mutant virus ensemble, a critical fact to design new antiviral strategies. We make the distinction between “intrinsic” and “extrinsic” properties of mutant distributions, and emphasize that there are several levels of complexity that can influence viral quasispecies behavior.

Keywords and phrases: RNA viruses, viral quasispecies, mutation rate, viral fitness, lethal mutagenesis

Mathematics Subject Classification: 92-02, 92B05, 92C40, 92D25

1. Introduction

Viruses are cellular parasites that have a genetic program encoded either in DNA or RNA and that requires a cell to be expressed. While the genome of DNA viruses ranges from a few thousand to several hundred thousand nucleotides, the genome of the RNA viruses identified to date is restricted to 3000-33000 nucleotides. That is, they have a very limited coding capacity to complete their replication cycles and to survive in a biosphere which is dominated by DNA-based unicellular and multicellular organisms. Some equilibrium has been reached between DNA and RNA genetic elements in life on Earth. While some models postulate an ancient cellular RNA world, no cells with an RNA genome have been discovered although, according to some views, RNA cells may still hide in some remote locations of our planet. This review article has as its main focus to describe and discuss some unique features of the population structure and dynamics of the RNA viruses that allow them to respond to selective constraints, including the host immune response. Viral RNA genomes can be single-stranded, double-stranded and with its

*Corresponding author. E-mail: edomingo@cbm.uam.es

information included in a single RNA molecule or in multiple RNA molecules (segmented genomes). The majority of viruses characterized to date are RNA viruses, and they include pathogens that will be familiar to all readers, such as the human immunodeficiency virus, the cause of the ongoing AIDS epidemic, hepatitis C virus, emerging viruses such as Dengue virus, the influenza virus that has affected most of us during our life times, and a number of animal and plant pathogens associated with economically devastating diseases. They have managed to survive for a long time by infecting relentlessly animals or plants, and by jumping from one host species to another.

The RNA viruses, together with other RNA genetic entities (retrovirus-like elements in mammalian genomes, the plant viroids, etc.) constitute one of several “RNA worlds” that have fascinated the scientific community since methods to approach life structures and mechanisms became available during the second half of the XXth century. Another RNA world is formed by multiple cellular RNAs that have been identified as playing essential regulatory roles in gene expression. Yet another is a likely RNA world that probably existed some four thousand million years ago, and that might have contained the most primitive replicating forms at the onset of life on Earth. Several general accounts of these different existing or hypothetical RNA worlds have been published, and the reader is referred to some books and articles as introductions to the subject [31, 41, 42, 51, 100].

Several fundamental questions have been asked: Why do RNA genetic elements in general, and RNA viruses in particular, exist? Are they remnants of a primitive RNA world or are they of a more recent origin? How do they manage to persist in a cellular DNA world? Do they play a function or are they simply “selfish” genetic entities? Have they contributed (and perhaps are still contributing) to cellular diversification and functional specilizations? Some of these questions have begun to be approached by comparing properties of the RNA genomes with those of the cells they infect. In this undertaking an interplay between theoretical modelling and experimental observation has been essential. In the last four decades, some unique features of RNA viruses have come to light: (i) High mutation rates during replication. (ii) A remarkable population heterogeneity and dynamics. Both traits resemble those predicted for primitive replicons during early life stages according to a theory of molecular evolution which is termed quasispecies theory. (iii) The occurrence of interactions among components of the heterogeneous viral populations that render mutant ensembles (rather than individual mutants) the true actors of virus behaviour. Here we examine the origin and scientific implications of these three features.

2. High mutation rates. Two births of quasispecies

Quasispecies is a theory of molecular evolution that formulated the mathematics of error-prone replication of primitive genetic elements. The initial, pioneer study was published by Manfred Eigen in 1971 [30]. It encompassed elements of information theory and Darwinian evolution, notably the concept of natural selection. The theory was intended to explain how primitive genetic elements (thought to be short RNA or RNA-like molecules) could self-organize to be able to display template activity (capacity to store and transmit information) and to achieve adaptability to ensure long-term propagation of the genetic message. An essential ingredient to accomplish self-organization and stability was the production of error copies in the right amounts: sufficient to ensure adaptability, but not in excessive numbers to avoid the genetic message losing its information contents. Quasispecies theory was subsequently further developed by Manfred Eigen and Peter Schuster (compiled in [32, 33]). The theory was ground breaking in that it represented the introduction of a molecular view into evolutionary biology. This was the birth of the theoretical quasispecies. Initially it was formulated as a deterministic theory, in mathematically solvable terms. Its key equation describes the production of error copies of a replicating entity that achieves a steady state mutant distribution in equilibrium (Figure 1). Extensions of quasispecies theory to non-equilibrium conditions (meaning finite genome populations in variable environments or fitness landscapes) have been developed in recent years [29, 77, 78, 98]. These extensions have rendered quasispecies theory more realistic to interpret the behaviour of real genetic entities, be putative primitive replicons or present day RNA viruses. In fact, there was a second birth of quasispecies from the experimental side.

Its description requires examination of pioneer studies with the most simple RNA viruses that have been described to date: the RNA bacteriophages.

In the early 1970's the RNA bacteriophages, a group of bacterial viruses with RNA as genetic material discovered in 1961 [57], became salient biological systems because they allowed for the first time the replication of a viral genome and the study of protein synthesis in the test tube. The only components required in the reaction mixture were the viral RNA as the template, a purified replicase, a host factor, nucleotide substrates, and adequate pH and ionic conditions. The viral RNA template (termed the plus strand) was copied into a complementary minus strand which in turn acted as template to produce more plus strands. The bacteriophage that best achieved in vitro replication had as its name $Q\beta$. A pioneer study by Sol Spiegelman and colleagues demonstrated Darwinian evolution (mutation, competition and selection) of $Q\beta$ RNA in vitro [61].

In his new Institute at the University of Zurich, Charles Weissmann the discoverer that replication of $Q\beta$ RNA occurred via a minus strand intermediate [35] adopted the minus strand-directed $Q\beta$ RNA synthesis in vitro as a system to introduce mutations at preselected positions of $Q\beta$ RNA. This was an important achievement that represented the origin of what is today known as reverse genetics [38,97]. It is remarkable that mutations could be introduced at preselected positions of the $Q\beta$ genome at a time when cDNA synthesis, in vitro DNA recombination, rapid nucleotide sequencing, and PCR technology were not available.

The second birth of quasispecies, came as a result of the study of one of the mutants of bacteriophage $Q\beta$ synthesised by site-directed mutagenesis that included a point mutation in the 3'- untranslated region [5,19,21]. Much to our surprise at the time, an extracistronic point mutant of bacteriophage $Q\beta$ was viable, meaning that it produced infectious progeny upon introduction of the RNA into *E.coli* spheroplasts (lysozyme-treated *E.coli* cells that were competent to uptake exogenous RNA). However, the mutant virus multiplied in *E.coli* at a rate slightly lower than that of the wild type virus. This was a lucky circumstance because it allowed quantification of the reversion of the mutant into the wild type, and to measure the selective advantage of the wild type relative to the mutant (a parameter that later became known as relative fitness in virology). A simple mathematical model written by Edward Batschelet from the Mathematics Department of the University of Zurich allowed an estimate of the rate at which the relevant reversion (a G- >A transition) occurred. The value was 10^{-4} mutational events per genome doubling [5]. This mutation rate was the first calculated for an RNA virus, and it was several orders of magnitude higher than the values that John Drake had previously obtained for DNA bacteriophages and bacteria [24]. Although the mutation rate referred to a precise mutation type at a single site of Q RNA, later studies with many RNA viruses using biochemical and genetic methods amply supported mutation rates averaging 10^{-3} to 10^{-5} substitutions per nucleotide copied (reviews in [25,79]). Growth-competition assays between two viruses (distinguishable genetically or phenotypically), carried out in cell culture or in a host organism, are currently used to quantify their relative fitness or replicative capacity [76] (Figure 2).

High mutation rates stand as a general feature of RNA viruses and other RNA genetic elements, and the biological consequences continue being unveiled at the time of this writing. When the rate of reversion of the $Q\beta$ mutant was obtained, it immediately provided a molecular interpretation of remarkable phenotypic heterogeneity and variation that characterized the RNA bacteriophages. For example, it had been observed that a class of mutants termed temperature-sensitive (ts) (because they cannot grow at high temperature) often arose in viral preparations that were supposed to have only wild type virus. Also, plaque size and morphology the appearance of the typical lesion of dead cells left by an infectious virus particle when applied onto a cell monolayer varied frequently within a viral population (as early examples and reviews, see [52,93,96]). It was soon recognized that the rate at which the extracistronic G- >A transition occurred (one event every 10.000 times that the replicase copied that specific template position) was unlikely to be exceptional. The reason was that passage of individual biological clones of phage $Q\beta$

even those derived by biological cloning of the wild type resulted in the generation of multiple mutants, in the absence of any mutagenic agent. The mutations mapped along the entire genome, suggesting high mutability of many sites in the RNA. The genetic heterogeneity generated as a result of a limited number of passages was such that each individual genome differed from the others in an average of 1 to 2 mutations per genome. As written in the publication where the generation of heterogeneity was described, “A Q β phage population is in a dynamic equilibrium, with viable genomes arising at a high rate on the one hand, and being strongly selected against on the other. The genome of Q β phage cannot be described as a defined, unique structure, but rather as a weighted average of a large number of different individual sequences” [21]. These dynamic features corresponded to those that quasispecies theory described for the hypothetical primitive replicons. At a meeting of the Max-Planck Society that took place in the Swiss village of Klosters in 1978, the connection between quasispecies theory of Eigen and Schuster and the experimental results of the Weissmann laboratory was made (reviews in [17,22,23]). Thus, the conclusion of these studies carried out four decades ago was that a real virus behaved as predicted by a theory of self-organization and adaptability of simple genetic elements. Remarkably, a concept developed to explain the behaviour of replicons that might have existed 4 thousand million years ago apparently was valid for a present day RNA virus. The Q β work represented the second birth of quasispecies, from experimental virology.

Evidence of high mutation rates and population heterogeneity was soon extended to animal and plant RNA viruses during the 1980's and early 1990's, establishing quasispecies dynamics as a general feature of animal and plant RNA viruses. ([As primary research references and reviews that reflect these early developments see [18,20,48–51,60,85]). Not a single exception, that is, an RNA virus devoid of a mutant spectrum, has been reported to date. During infections, even those generated by a single particle (a biological cloning event), mutant distributions (also referred to as mutant spectra or mutant clouds) are rapidly generated. A term used frequently to describe the basic features of RNA genetic elements is “error-prone replication”. Despite mistakes in the copying of messages (genetic or other) being universal, some biochemical features of the DNA and RNA replication machineries can be identified as determining their extent.

3. Molecular basis of error-prone replication

The biochemical basis of the high mutation rates displayed by RNA viruses is the absence of a 3'-to-5' exonuclease in their RNA-dependent RNA polymerases (RdRps, often referred to as RNA replicases, the enzymes responsible for RNA genome replication). This exonuclease acts as a proofreading-repair function in the cellular replicative DNA-dependent DNA polymerases by removing incorrectly incorporated nucleotides at the growing 3' end of the DNA [39]. No evidence of a proofreading-repair activity was obtained using several biochemical assays with vesicular stomatitis virus (VSV) polymerase *in vitro* [variation of pyrophosphate levels, or of nucleoside-triphosphate concentrations, incorporation of nucleoside [1-thio]-triphosphate analogues] [88]. These negative results agree with the absence in viral RdRps of a domain corresponding to a 3'-to-5' exonuclease activity, as evidenced by sequence alignments and by the comparison of three-dimensional structures of DNA polymerases and RNA polymerases [6,37,87,89].

A significant exception to the absence of a proofreading-repair function in RNA viruses is the presence of a functional 3'-to-5' exonuclease domain in a non-structural protein (nsp 14) of coronaviruses [28,62], the viruses with the largest unsegmented RNA genomes. The coronaviruses in which the exonuclease has been rendered non-functional by mutation display a 12- to 15-fold decrease in template-copying accuracy relative to the wild type virus, as evidenced by an increase of the mutant spectrum complexity [15,27]. Other related viruses termed Nidovirales which have a smaller genome than the coronaviruses do not have the exonuclease function. The presence of a functional proofreading-repair enzyme in the largest RNA viruses known is consistent with the prediction that a large amount of genetic information to be transmitted demands a correspondingly higher accuracy of the replication machinery. This was expressed in quasispecies theory in the form of an error threshold relationship [33] that has found a practical application in virology as a new antiviral strategy termed lethal mutagenesis, consisting in the extinction

of virus through introduction of an excess of mutations in their genomes during replication (discussed in a later section).

RdRp fidelity is not dependent only on the presence or absence of a proofreading-repair activity because it can be modified as a result of amino acid substitutions that directly or indirectly affect the interaction of the polymerase-template-primer complex with the incoming nucleotide [36]. This has been documented with a variety of high and low fidelity mutants of several RNA viruses, an ongoing line of active research [11, 22, 59, 94]. High fidelity mutants give rise to mutant spectra of lower complexity (a lower average number of mutations per genome in the mutant swarm) that often show decreased capacity to adapt to complex environments, relative to the corresponding virus that displays a standard mutation rate and higher mutant spectrum complexity [11, 75, 94, 95]. The behavior of viral mutants with modified polymerase fidelity constitutes a direct proof of the selective advantage of high mutation rates for viruses, and a demonstration of the important biological role played by mutant spectra [7, 22] (Figure 1).

An additional mechanism that contributes to keeping misincorporations introduced in RNA is the inability of post-replicative repair pathways present in the cell to act on replicative RNA forms. The cells have evolved several enzymes and accessory proteins whose function is to repair damage in DNA or to correct mispaired bases in double stranded DNA when the misincorporation event went unnoticed by the polymerase proofreading activity [39, 56, 64]. Again, biochemical mechanisms fit evolutionary predictions in the sense that a large DNA genome such as the human genome which has a total number of nucleotide pairs which is 10^5 - to 10^6 -fold larger than the number of nucleotides in RNA virus genomes must restrict the number of mutations due to the need to preserve a complex inheritable information [7, 33, 51]. In other words, the negative effects of a given number of mutations per nucleotides are amplified in complex genomes.

4. Influence of the quasispecies concept in the understanding of RNA virus behavior

The recognition that replicating populations of RNA viruses consist of mutant clouds rather than defined genomic sequences (Figure 1) provided an interpretation of their broad adaptability, and it represented a rupture with concepts of classic population genetics. The main reason of the rupture is that mutation was considered a rare and often deleterious event (see for example several statements in [99] discussed in [73]). In contrast, a tenet of quasispecies is that mutation plays a critical role in adaptation. In fact, of the several mathematical formulations of evolutionary dynamics, which are all interconnected, quasispecies has as its distinctive characteristic its emphasis on mutation [67]. In the RNA viruses, mutation can no longer be considered an occasional event. Rather, mutation is incessant and a fundamental part of their natural life cycles.

Opposition to viral quasispecies expressed by some population geneticists came in the form of incredulity in high mutation rates. Some attributed the viral population heterogeneities partly to copying mistakes in the enzymatic reactions involved in reverse transcription and polymerase chain reaction amplification (RT-PCR). It was also expressed in manifestos in favor of population genetics being sufficient to explain RNA virus evolution, without the need of quasispecies theory entering the picture. Furthermore, the deterministic nature of the initial quasispecies formulation was also taken as an argument against the theory being adequate for RNA viruses. These controversies have been recently commented [73] and somehow reaffirm that quasispecies has indeed introduced a new point of view in our understanding of RNA viruses. Had the quasispecies irruption represented a trivial addition, no opposition would have been manifested. Quasispecies has influenced our perception not only of virus evolution but also of viral pathogenesis. This is a key point that has established an interesting link between what one could call basic population virology and clinical virology. The ultimate reason of why viral quasispecies is of relevance to viral disease is that one or very few mutations introduced in a viral genome (in the mutant swarms such as those depicted in Figure 1) can profoundly alter the response of viruses to frequently encountered selective constraints, including the host immune response or externally administered drugs. In particular,

mutation frequencies are such that inhibitor-resistant mutants of human immunodeficiency virus type 1 (HIV-1), hepatitis B virus (HBV) and hepatitis C (HCV) are often present in the mutant spectra of virus present in patients that had never been treated with the corresponding inhibitor (reviewed in [22]). Implications of genetic variation in virus adaptability are addressed next.

5. The phenotypic effects of single mutations and of more drastic genetic variation

Mutation is the most widespread type of genetic change. However, RNA viruses, as any other type of cellular or subcellular genetic entity, can undergo more drastic genetic change (Figure 3). Recombination is the process by which two parental genomes give rise to a chimeric genome that includes one or more regions from one parent and one or more regions from the other parent. Recombination occurs in DNA and RNA genomes, although in the case of RNA viruses is more frequent among the positive-strand RNA viruses (those in which genomic RNA and viral mRNAs have the same polarity) than among the negative-strand RNA viruses (those in which genomic RNA and all or part of viral mRNAs have different polarity) [47]. RNA recombinants can arise when an RdRp which is in the process of copying a template RNA molecule jumps at the equivalent position of a second template molecule. Since no discontinuity in phosphodiester bond formation occurs, the result is the formation of a chimeric genome with the crossing point at a site where the sequences are identical or nearly-identical. This type of recombination is termed homologous recombination, and it requires RNA replication.

Recombination can also be non-homologous and non-replicative, and it can link together RNA pieces of very different origin, including viral RNAs with cellular RNA (Figure 3) (reviewed in [1, 82, 91]). Recombination in virus evolution can serve at least two different purposes: to explore new combinations of genetic material for adaptation to new environments, or to rescue viable genomes from parental genomes that had acquired an excessive mutational load. In this sense, recombination serves as a fitness-rescuing mechanism. Nucleotide sequence comparisons have provided evidence that some present day viruses probably arose by recombination. A classical example is the alphavirus Western equine encephalitis virus which was probably the result of a recombination event between two viruses, one resembling Sindbis virus and the other resembling Eastern equine encephalitis virus. A large proportion of pathogenic polioviruses that arise in world areas as a cause of poliomyelitis consist of recombinants between poliovirus vaccine strains and human enteroviruses [1]. Thus, recombination is a powerful influence in the emergence of new viral pathogens. Evidence that past recombination events have shaped current versions of circulating viruses has been obtained with salient pathogens such as hepatitis B virus (a virus that while having a DNA genome, its replication cycle includes an error-prone reverse transcription step) [83].

While mutation rates and frequencies for RNA and DNA genomes have at least some range of consistent values, it has not been possible to obtain a unified view on recombination rates and frequencies, probably due to their inherent variability and also to technical difficulties for their measurement. For some picornaviruses and coronaviruses, the frequency of recombinants can reach 10% to 20% of the total progeny in a typical infection, and the recombination rate of HIV-1 is about 5-fold larger than the mutation rate [80]. Limited polymerase processivity (the capacity of the enzyme to continue copying the same template molecule) appears to be one of the factors that determine high recombination rates by template jumping. It is likely that many recombination events occur during viral infections but they go unnoticed because of the lack of suitable markers and absence of a driving force to confer an advantage to the recombinants. Under selective drug pressure, dually or multiply drug-resistant HIV-1 recombinants can be selected [9, 55, 63].

Gains or losses of nucleotide stretches are termed insertions and deletions, respectively, often abbreviated as “indels”. They occur by a variety of recombination mechanism either within a template molecule (intra-molecular event) or with the participation of two (on occasions perhaps more) template molecules

(inter-molecular event). While prior mutations in the parental genomes are a necessity for homologous recombination to have biological significance, “indels” can be generated by non-homologous recombination involving identical (non-mutated) genomes. Other genetic modifications such as inversions, translocations, fusions and fissions are rare in viruses but have been described in viral-like cellular elements and in cellular DNA, including cancer cells.

An interesting observation that applies to viruses as well as to complex cellular genomes is that the extent of a genetic lesion does not necessarily predict the magnitude of a phenotypic consequence. A virus may tolerate indels at several genomic sites with no or little apparent effect while a point mutation that affects directly or indirectly the active site of a viral enzyme can be lethal. In the human genome some points mutations may be the cause of more severe genetic disease than drastic chromosome inversions.

6. Internal interactions within quasispecies. How mutant ensembles can influence therapeutic outcomes

The behavior of a viral quasispecies cannot be predicted from the behavior of its individual components in isolation, fundamentally due to internal interactions which are established among individual components of a mutant spectrum. Complementing interactions of a positive nature ([manifested because fitness of mutant ensembles is higher than fitness of individual biological clones from the same ensemble [21, 26]) may enhance the replicative capacity of a quasispecies. In contrast, interfering interactions, mediated by intra-population dominant-negative mutants, may suppress replication by deterioration of viral functions [12, 43, 74] (Figure 4). For an antiviral approach to be effective, the presence of a mutant spectrum must be taken into consideration not only because the spectrum may include mutants with potential for adaptation to the antiviral drugs, but also because intra-mutant spectrum interactions may modify the response to the drugs. Learning how to relate viral population composition to the response to drugs may avoid or delay selection of drug-resistant mutants and the ensuing treatment failure.

The existence of an error threshold for maintenance of genetic information [29, 33, 65] (see section on “Molecular basis of error-prone replication”) launched investigations on a new antiviral strategy termed lethal mutagenesis. It consists on increasing the error rate during viral replication in the presence of a virus-specific mutagenic agent. Current evidence suggests that virus extinction through enhanced mutagenesis may involve a shift from dominance of complementation among components of the mutant spectrum, to a dominance of interfering interactions as a result of progressively higher concentration of mutants expressing altered proteins (Figure 4). As the mutagenic activity continues, the frequency of genomes harboring multiple mutations and lethal combinations of mutations increases, contributing eventually to the replicative collapse of the ensemble. This proposal agrees with the observation that upon subjecting a replicating viral population to enhanced mutagenesis, the decrease in viral infectivity preceded the decrease in viral RNA levels [44]. These results suggest the presence of a class of RNA replication-competent (or partly competent) genomes that, although partially or completely incapable of completing an infectious cycle on their own, can in some way perturb functions of the standard (fully infectious) genomes coexisting in the same quasispecies [44, 53].

A transition towards dominance of interfering interactions has been further supported by studies with foot-and-mouth disease virus (FMDV) both, by interference exerted by heavily mutagenized, preextinction FMDV populations [4], and by interference exerted by specific mutants that express altered viral proteins [74]. It has been postulated that negative-dominant proteins could contribute to the suppression of replication of standard virus, because of the formation of inactive protein complexes. Negative-dominant poliovirus (PV) proteins were shown to dominantly interfere with the growth of drug-resistant PV mutants [12]. These observations match the concept that a viral population as a whole can act as the unit of selection. Because an RNA virus population is made of a spectrum of viable mutants that display a range of different fitness values [21, 26], it is realistic to force negative intra-population interactions to

produce a collective fitness decrease of the entire quasispecies. To aim at a fitness decrease may be as valid an antiviral strategy as to aim at a reduction of viral load [10,22].

An effective suppression of viral infectivity may have a dual consequence of reducing the viral load and achieving viral extinction. Reductions of viral load need not imply virus extinction but may result in a very low replication level that can facilitate clearance of the virus by the immune system. Quantified decreases of viral load, obtained through non-mutagenic inhibitors, did not result in viral extinction, while equal decreases of viral load, achieved in parallel with a mutagenic agent, lead to extinction [2,45,68–70,81]. These experimental results highlight a critical role of mutagenesis (not only its effect in decreasing viral load) in extinction. A decrease of the basic reproductive ratio (R_0) to values that impede a sustained cell-to-cell spread of a virus is not the direct cause of virus extinction, since the decrease is a consequence of mutagenesis (R_0 is used in epidemiology but here we apply it to intra-host virus replication). The essential contribution of enhanced mutagenesis to viral extinction has been further underlined by the observation that drastic reductions of population size (such as those involved in plaque-to-plaque transfers) acting on a virus debilitated by accumulation of mutations, are not sufficient to extinguish the population [34].

To achieve sustained decreases of viral load will require an understanding of the molecular mechanisms that underlie lethal mutagenesis. Only with such a knowledge adequate treatment protocols that involve mutagens and inhibitors administered sequentially or in combination, can be properly designed, as indicated by recent results. In the transition towards loss of infectivity mediated by lethal mutagenesis associated with ribavirin treatment, we observed a movement in sequence space towards hypermutated genomes enriched in A, U nucleotides that included defective but replication-competent genomes [72]. A, U-biased mutants accumulated in the mutant spectrum as a result of ribavirin mutagenesis, and the altered codon triplets contributed to the transition towards extinction. Interfering mutants generated by a mutagen can not replicate in the presence of an inhibitor of viral replication and, as a result, their interfering activity that participates in extinction is prevented. The interaction between a mutagenic agent and an inhibitor regarding their joint effect on a viral population was investigated with capsid and polymerase mutants of FMDV that had been characterized biochemically and structurally [4,58,74]. Following coelectroporation of cells with the relevant RNAs, an excess of replication-competent mutants caused a strong and specific interference on FMDV replication. While the presence of the nucleoside analogue ribavirin during viral replication allowed interference to occur, the presence of the inhibitor of FMDV replication guanidine hydrochloride (GU) prevented the interfering activity [71]. We interpreted this result as evidence that the interfering genomes must be allowed to replicate to exert their activity. This is also consistent with the observation that a double polymerase FMDV mutant that rendered an RNA replication-competent virus lost its interfering activity when a third polymerase mutation that abolished RNA replication was introduced [74]. Furthermore, a model of FMDV population dynamics predicted that, independently of the need of defectors to be RNA synthesis positive, the simultaneous presence of a mutagen and an inhibitor may decrease the probability of extinction because the mutagen may rise the level of inhibitor-resistant mutants in the population [54,71].

Thus, despite the amply recognized advantage of combination therapies over monotherapy to avoid or delay selection of escape mutants when only non-mutagenic inhibitors are involved in therapy, a sequential inhibitor-mutagen administration can have an advantage when a mutagenic agent is a component of therapy. In the experiments with FMDV, the opposite sequential administration of ribavirin first, followed by the inhibitor GU, resulted in a faster selection of inhibitor-escape mutants than when the same initial viral load (lacking the previous ribavirin treatment) was subjected to GU alone. A sufficiently low viral load, achieved as a consequence of a mutagenic treatment, may allow continued, low-level viral replication without selection of inhibitor-escape mutants despite the presence of the inhibitor [70]. The range of parameters (mainly those relating intensity of inhibition and mutagenesis) that may favour the effectiveness of a sequential versus combination inhibitor-mutagen treatments have been investigated using a theoretical model that mimics *in silico* the interaction between the inhibitor and the mutagenic agent [54]. Experimental and theoretical studies on the adequacy of sequential versus combination treatments continue at the moment, using additional viruses that display different replication mechanisms.

7. Several levels of complexity in virology

The trajectory from quasispecies theory to the new prospects of antiviral interventions summarized in previous paragraphs is a demonstration that the behavior of viral populations can hardly be predicted from the behavior of the individual viral genomes that compose it. Although not as fashionable as a decade ago, the term “complexity” is used to refer to systems composed of interconnected parts, in which the performance of the system as a whole is not obvious from the performance of the individual parts. Expressed in more general terms, the sciences of complexity study those systems in which there is no simple and predictable relationship between levels, between the properties of parts and of wholes [86]. The brain and the immune system behave as complex, adaptive systems.

Regarding viral quasispecies, the term complexity has had two different meanings: (i) the amount of information (in regulatory, non-coding and protein-coding regions) present in each RNA genomic molecule (in each of the horizontal lines in Figure 1A), and (ii) the variations in nucleotide sequence considering a collection of genomes such as those present in a mutant spectrum (the ensemble of horizontal lines in any of the distributions shown in Figure 1A). For example, we can refer to a mutant spectrum with an average mutation frequency of 10^{-4} substitutions per nucleotide as having a 10-fold higher complexity than a spectrum with an average mutation frequency of 10^{-5} substitutions per nucleotide. It is this second meaning of complexity the one that bears a direct relationship with the concept of complexity originated in physics, and that has ramifications in multiple fields of science such as sociology, economics, informatics or biology [see [40,84] for an introduction to the general concept of complexity] and [86] for its implications in biology.

In virology an interesting consideration is the connection between population size and complexity. If we take mutation frequency as a quantification of complexity, then the population size is not a relevant parameter (except for a trivial inaccuracy of the calculation due to a limitation in sample size). That is, 10 mutations in a total of 10,000 nucleotides sequenced (1000 nucleotides in each of a total population of 10 genomes) gives the same mutation frequency as 100 mutations in a total of 100,000 nucleotides sequenced (1000 nucleotides in each of a total population of 100 genomes): 10^{-3} substitutions per nucleotide (Figure 5). Using terminology from physics, the mutation frequency is an “intrinsic” property, independent of the population size. However, if we consider the presence of specific and phenotypically important variants in the mutant spectrum, the property of the mutant spectrum becomes “extrinsic”. That is, a specific mutant present at a frequency of 10^{-3} (a reasonable frequency for some mutant types) is far less likely to be present in a population of 10 genomes than of 1000 genomes. Thus, in this respect, the behavior (in this case adaptive capacity) of a virus is strongly dependent on the population size. Although only indirectly related to complexity, the “extrinsic” (population size-dependent) properties of mutant spectra contribute to their unpredictable behavior.

A different manifestation of complexity is that the mutant spectrum, depending on its size and composition, may affect the fate (increase or decrease in frequency) of individual variants. This was predicted by quasispecies theory [90] and demonstrated experimentally first by de la Torre and Holland working with VSV [14]. That some types of viral mutants can be suppressed by mutant spectra has been confirmed in cell culture and in vivo [8,13,66,92]. Thus, it is uncertain whether a specific variant immersed in a mutant spectrum may reach dominance, despite the variant displaying high fitness relative to the surrounding spectrum (reviewed in [22]). The population size-dependent suppression of specific variants by a mutant spectrum, irrespective of the complexity of the mutant spectrum, constitutes a demonstration of the potential biological consequences of the extrinsic properties of viral quasispecies.

Virology embodies additional levels of complexity that lie beyond the events that occur at the viral intrapopulation level. The major ones relate to the process of emergence and re-emergence of viral disease, a salient problem in public health and in agriculture. This timely topic has been amply reviewed and popularized [3,16,46], and will be briefly discussed here in connection with complexity. True viral emergences (the appearance of a new pathogen not previously described for the affected host species) or

re-emergences (the re-appearance of a known pathogen after a silent period or in a new geographical area) are a consequence of a series of environmental, ecologic, and sociologic influences. Such interconnected factors allow a virus present in our biosphere to come into contact with a potentially new host. It has not been possible to explain a viral emergence (i.e. HIV-1, SARS coronavirus, Schmallenberg virus, etc.) by a single influence or even by a sum of influences. Given the number of interactions between climate, flora, fauna, distribution of viral vectors and potential hosts, etc. viral emergences are the unpredictable outcome of multiple properties of the involved parts. Viral diseases have emerged in the human populations at a rate of one per year during the last decades. Therefore, it is highly likely that they will continue to emerge, but we cannot know when and where they will occur. Thus, this major concern of public health belongs also to complexity, at a second level that follows the interactions within mutant spectra of viral populations. Interestingly, both levels are not independent. Most emerging viral infections are of a zoonotic origin. They originated from some animal reservoir, and in each animal the virus replicates as multiple quasispecies. It is a matter of chance what types of variants will be present in (or shed by) the animal, and that may come in contact with a potential new host, human or other. Thus, chance events occurring at the level of mutant spectra are interwoven with the chance events at the environmental and sociological level. Complexity has truly permeated virology.

8. Concluding remarks

RNA viruses, and any genetic elements whose replication is catalyzed by low fidelity enzymes (or enzyme complexes) will generate mutant swarms as they multiply in their host species. Because regulatory signals and expression products in the genomes that populate our biosphere exert activities in cis (on the same genome that harbors or expresses them) and in trans (on genomes other than the one that harbors or expresses them), interactions are unavoidably exerted among components of the same swarm. There are several manifestations of the collective behaviour of viral quasispecies (reviewed in [22]), and here we have emphasized suppression of individual variants by the entire mutant ensembles where they are immersed. The fact that the molecular interactions among genomes, their expression products and host factors are multiple and poorly understood, renders viral quasispecies performance within the scope of biological complexity. It is now well established both, through studies of experimental evolution in the laboratory and in clinical practice, that viral populations are designed to respond to selective constraints. They are reactive, adaptive and contextual. Significantly, these are three attributes employed in modern neurology to refer to the activities of the brain, a paradigm of biological complexity.

Generation of mutant swarms is the first stage in virus diversification in nature. Genetic and phenotypic diversification of viruses has been extensively documented both, in designed laboratory experiments (with cells in culture and host organisms) and by analysis of natural viral isolates. Molecular epidemiology has taught us that most RNA viruses evolve at rates of 10^{-3} to 10^{-4} mutations per nucleotide and year of evolutionary time. Virus spread in nature, and the encounter of potentially new host species (such as when a simian immunodeficiency virus met a human host to become established as human immunodeficiency virus type 1, the cause of AIDS) are influenced by multiple interconnected factors. Indeed, infected animal hosts and vectors move as a result of sociological and environmental (climatic) influences. They are interconnected: deforestation in the Amazonian region to raise cattle may increase tick populations that may come in contact with migrating animals that modify their routes due to climatic change, and in each host harbouring a replicating virus there is the chance generation of specific variants, etc., etc. To the point that we can develop predictive tools for complex systems we may be able to design new control strategies for viral diseases while learning basic principles of complex behavior in the fascinating world of viruses.

Acknowledgements. Work in the authors' laboratory supported by grants BFU2011-23604 from Ministerio de Economía y Competitividad, 360766/09 from FIPSE and IDI-20110115 from CDTI. CIBERehd is funded by Instituto de Salud Carlos III. C.P. has a contract from CIBERehd.

References

- [1] V.I. Agol. *Picornaviruses as a model for studying the nature of RNA recombination*. In: The Picornaviruses. E. Ehrenfeld, E. Domingo and R.P. Roos, eds, ASM Press, Washington DC. pp. 239–252, 2010.
- [2] A. Airaksinen, N. Pariente, L. Menendez-Arias, E. Domingo. *Curing of foot-and-mouth disease virus from persistently infected cells by ribavirin involves enhanced mutagenesis*. Virology, 311 (2003), 339–349.
- [3] R. Antia, R.R. Regoes, J.C. Koella, C.T. Bergstrom. *The role of evolution in the emergence of infectious diseases*. Nature, 426 (2003), 658–661.
- [4] A. Arias, R. Agudo, C. Ferrer-Orta, R. Perez-Luque, A. Airaksinen, E. Brocchi, E. Domingo, N. Verdagner, C. Escarmis. *Mutant viral polymerase in the transition of virus to error catastrophe identifies a critical site for RNA binding*. J. Mol. Biol., 353 (2005), 1021–1032.
- [5] E. Batschelet, E. Domingo, C. Weissmann. *The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate*. Gene, 1 (1976), 27–32.
- [6] A. Bernad, L. Blanco, J.M. Lazaro, G. Martin, M. Salas. *A conserved 3' 5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases*. Cell, 59 (1989), 219–228.
- [7] C.K. Biebricher, M. Eigen. *The error threshold*. Virus Res., 107 (2005), 117–127.
- [8] B. Borrego, I.S. Novella, E. Giralt, D. Andreu, E. Domingo. *Distinct repertoire of antigenic variants of foot-and-mouth disease virus in the presence or absence of immune selection*. J. Virol., 67 (1993), 6071–6079.
- [9] M.J. Buzon, T. Wrin, F.M. Codoner, J. Dalmau, P. Phung, A. Bonjoch, E. Coakley, B. Clotet, J. Martinez-Picado. *Combined antiretroviral therapy and immune pressure lead to in vivo HIV-1 recombination with ancestral viral genomes*. Journal of acquired immune deficiency syndromes (1999), 57 (2011), 109–117.
- [10] M. Clementi. *Perspectives and opportunities for novel antiviral treatments targeting virus fitness*. Clin Microbiol Infect, 14 (2008), 629–631.
- [11] L.L. Coffey, Y. Beeharry, A.V. Borderia, H. Blanc, M. Vignuzzi. *Arbovirus high fidelity variant loses fitness in mosquitoes and mice*. Proc Natl Acad Sci U S A, 108 (2011), 16038–16043.
- [12] S. Crowder, K. Kirkegaard. *Trans-dominant inhibition of RNA viral replication can slow growth of drug-resistant viruses*. Nature Genetics, 37 (2005), 701–709.
- [13] K.M. Chumakov, L.B. Powers, K.E. Noonan, I.B. Roninson, I.S. Levenbook. *Correlation between amount of virus with altered nucleotide sequence and the monkey test for acceptability of oral poliovirus vaccine*. Proc. Natl. Acad. Sci. USA, 88 (1991), 199–203.
- [14] J.C. de la Torre, J.J. Holland. *RNA virus quasispecies populations can suppress vastly superior mutant progeny*. J. Virol., 64 (1990), 6278–6281.
- [15] M.R. Denison, R.L. Graham, E.F. Donaldson, L.D. Eckerle, R.S. Baric. *Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity*. RNA biology, 8 (2011), 270–279.
- [16] E. Domingo. *Mechanisms of viral emergence*. Vet Res, 41 (2010), 38.
- [17] E. Domingo, C. Biebricher, M. Eigen, J.J. Holland. *Quasispecies and RNA Virus Evolution: Principles and Consequences*. Landes Bioscience, Austin, 2001.
- [18] E. Domingo, M. Davila, J. Ortin. *Nucleotide sequence heterogeneity of the RNA from a natural population of foot-and-mouth-disease virus*. Gene, 11 (1980), 333–346.
- [19] E. Domingo, R.A. Flavell, C. Weissmann. *In vitro site-directed mutagenesis: generation and properties of an infectious extracistronic mutant of bacteriophage Q*. Gene, 1 (1976), 3–25.
- [20] E. Domingo, J.J. Holland, P. Ahlquist. *RNA Genetics*, CRC Press, Boca Raton, 1988.
- [21] E. Domingo, D. Sabo, T. Taniguchi, C. Weissmann. *Nucleotide sequence heterogeneity of an RNA phage population*. Cell, 13 (1978), 735–744.
- [22] E. Domingo, J. Sheldon, C. Perales. *Viral quasispecies evolution*. Microbiology and Molecular Biology Reviews, 76 (2012), 159–216.
- [23] E. Domingo, S. Wain-Hobson. *The 30th anniversary of quasispecies. Meeting on 'Quasispecies: past, present and future'*. EMBO Rep, 10 (2009), 444–448.
- [24] J.W. Drake. *Comparative rates of spontaneous mutation*. Nature, 221 (1969), 1132.
- [25] J.W. Drake, J.J. Holland. *Mutation rates among RNA viruses*. Proc. Natl. Acad. Sci. USA, 96 (1999), 13910–13913.
- [26] E.A. Duarte, I.S. Novella, S. Ledesma, D.K. Clarke, A. Moya, S.F. Elena, E. Domingo, J.J. Holland. *Subclonal components of consensus fitness in an RNA virus clone*. J. Virol., 68 (1994), 4295–4301.
- [27] L.D. Eckerle, M.M. Becker, R.A. Halpin, K. Li, E. Venter, X. Lu, S. Scherbakova, R.L. Graham, R.S. Baric, T.B. Stockwell, D.J. Spiro, M.R. Denison. *Infidelity of SARS-CoV Nsp14-exonuclease mutant virus replication is revealed by complete genome sequencing*. PLoS Pathog, 6 (2010), e1000896.
- [28] L.D. Eckerle, X. Lu, S.M. Sperry, L. Choi, M.R. Denison. *High fidelity of murine hepatitis virus replication is decreased in nsp14 exonuclease mutants*. J Virol, 81 (2007), 12135–12144.
- [29] M. Eigen. *Natural selection: a phase transition?* Biophys. Chem., 85 (2000), 101–123.
- [30] M. Eigen. *Self-organization of matter and the evolution of biological macromolecules*. Naturwissenschaften, 58 (1971), 465–523.
- [31] M. Eigen. *Steps towards life*, Oxford University Press, 1992.
- [32] M. Eigen, J. McCaskill, P. Schuster. *Molecular quasi-species*. J. Phys. Chem., 92 (1988), 6881–6891.
- [33] M. Eigen, P. Schuster. *The hypercycle. A principle of natural self-organization*. Springer, Berlin, 1979.

- [34] C. Escarmis, E. Lazaro, A. Arias, E. Domingo. *Repeated bottleneck transfers can lead to non-cytocidal forms of a cytopathic virus: implications for viral extinction*. J. Mol. Biol., 376 (2008), 367–379.
- [35] G. Feix, R. Pollet, C. Weissmann. *Replication of viral RNA, XVI. Enzymatic synthesis of infectious viral RNA with noninfectious Q-beta minus strands as template*. Proc Natl Acad Sci U S A, 59 (1968), 145–152.
- [36] C. Ferrer-Orta, R. Agudo, E. Domingo, N. Verdaguer. *Structural insights into replication initiation and elongation processes by the FMDV RNA-dependent RNA polymerase*. Current Opinion in Structural Biology, 19 (2009), 752–758.
- [37] C. Ferrer-Orta, A. Arias, R. Agudo, R. Perez-Luque, C. Escarmis, E. Domingo, N. Verdaguer. *The structure of a protein primer-polymerase complex in the initiation of genome replication*. EMBO J, 25 (2006), 880–888.
- [38] R.A. Flavell, D.L. Sabo, E.F. Bandle, C. Weissmann. *Site-directed mutagenesis: generation of an extracistronic mutation in bacteriophage Q beta RNA*. J Mol Biol, 89 (1974), 255–272.
- [39] E.C. Friedberg, G.C. Walker, W. Siede, R.D. Wood, R.A. Schultz, T. Ellenberger. *DNA repair and mutagenesis*. American Society for Microbiology, Washington, DC, 2006.
- [40] M. Gell-Mann. *Complex adaptive systems*, in: G.A. Cowan, D. Pines, D. Meltzer (Eds.). Complexity. Metaphors, models and reality, Wesley Publishing Co., Reading, MA, 1994, pp. 17–45.
- [41] R.F. Gesteland, T.R. Cech, J.F. Atkins. *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2006.
- [42] W. Gilbert. *The RNA world*. Nature, 319 (1986), 618.
- [43] C. Gonzalez-Lopez, A. Arias, N. Pariente, G. Gomez-Mariano, E. Domingo. *Preextinction viral RNA can interfere with infectivity*. J. Virol., 78 (2004), 3319–3324.
- [44] A. Grande-Perez, E. Lazaro, P. Lowenstein, E. Domingo, S.C. Manrubia. *Suppression of viral infectivity through lethal defection*. Proc. Natl. Acad. Sci. USA, 102 (2005), 4448–4452.
- [45] A. Grande-Perez, S. Sierra, M.G. Castro, E. Domingo, P.R. Lowenstein. *Molecular indetermination in the transition to error catastrophe: systematic elimination of lymphocytic choriomeningitis virus through mutagenesis does not correlate linearly with large increases in mutant spectrum complexity*. Proc. Natl. Acad. Sci. USA, 99 (2002), 12938–12943.
- [46] B.L. Haagmans, A.C. Andeweg, A.D. Osterhaus. *The application of genomics to emerging zoonotic viral diseases*. PLoS Pathog, 5 (2009), e1000557.
- [47] G.Z. Han, M. Worobey. *Homologous recombination in negative sense RNA viruses*. Viruses, 3 (2011), 1358–1373.
- [48] J.J. Holland. *Continuum of change in RNA virus genomes*. In: A.L. Notkins, M.B.A. Oldstone (Eds.). Concepts in Viral Pathogenesis, Springer-Verlag, New York, 1984.
- [49] J.J. Holland. *Genetic diversity of RNA viruses*. Current Topics in Microbiology and Immunology, Springer-Verlag, Berlin, 1992.
- [50] J.J. Holland, E.A. Grabau, C.L. Jones, B.L. Semler. *Evolution of multiple genome mutations during long-term persistent infection by vesicular stomatitis virus*. Cell, 16 (1979), 495–504.
- [51] J.J. Holland, K. Spindler, F. Horodyski, E. Grabau, S. Nichol, S. VandePol. *Rapid evolution of RNA genomes*. Science, 215 (1982), 1577–1585.
- [52] K. Horiuchi. *Genetic studies of RNA phages*, in: N.D. Zinder (Ed.) RNA Phages, Cold Spring Harbor laboratory, Cold Spring Harbor, New York, 1975, pp. 29–50.
- [53] J. Iranzo, S.C. Manrubia. *Stochastic extinction of viral infectivity through the action of defectors*. Europhys. Lett., 85 (2009), 18001.
- [54] J. Iranzo, C. Perales, E. Domingo, S.C. Manrubia. *Tempo and mode of inhibitor-mutagen antiviral therapies: A multidisciplinary approach*. Proc Natl Acad Sci U S A, 108 (2011), 16008–16013.
- [55] K.S. Kemal, C.M. Kitchen, H. Burger, B. Foley, D. Mayers, T. Klimkait, F. Hamy, K. Anastos, K. Petrovic, V.N. Minin, M.A. Suchard, B. Weiser. *Recombination Between Variants from Genital Tract and Plasma: Evolution of Multidrug-Resistant HIV Type 1*. AIDS research and human retroviruses, (2012), in press.
- [56] G.M. Li. *Mechanisms and functions of DNA mismatch repair*. Cell research, 18 (2008), 85–98.
- [57] T. Loeb, N.D. Zinder. *A bacteriophage containing RNA*. Proc Natl Acad Sci U S A, 47 (1961), 282–289.
- [58] R. Mateo, A. Diaz, E. Baranowski, M.G. Mateu. *Complete alanine scanning of intersubunit interfaces in a foot-and-mouth disease virus capsid reveals critical contributions of many side chains to particle stability and viral function*. J Biol Chem, 278 (2003), 41019–41027.
- [59] L. Menendez-Arias. *Mutation rates and intrinsic fidelity of retroviral reverse transcriptases*. Viruses, 1 (2009), 1137–1165.
- [60] A. Meyerhans, R. Cheynier, J. Albert, M. Seth, S. Kwok, J. Sninsky, L. Morfeldt-Manson, B. Asjo, S. Wain-Hobson. *Temporal fluctuations in HIV quasispecies in vivo are not reflected by sequential HIV isolations*. Cell, 58 (1989), 901–910.
- [61] D.R. Mills, R.L. Peterson, S. Spiegelman. *An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule*. Proc. Natl. Acad. Sci. USA, 58 (1967), 217–224.
- [62] E. Minskaia, T. Hertzog, A.E. Gorbalenya, V. Campanacci, C. Cambillau, B. Canard, J. Ziebuhr. *Discovery of an RNA virus 3'–>5' exoribonuclease that is critically involved in coronavirus RNA synthesis*. Proc. Natl. Acad. Sci. USA, 103 (2006), 5108–5113.
- [63] L. Moutouh, J. Corbeil, D.D. Richman. *Recombination leads to the rapid emergence of HIV-1 dually resistant mutants under selective drug pressure*. Proc Natl Acad Sci U S A, 93 (1996), 6106–6111.
- [64] H. Naegeli. *Mechanisms of DNA damage recognition in mammalian cells*. Landes Bioscience, Austin, Texas, 1997.
- [65] M.A. Nowak, P. Schuster. *Error thresholds of replication in finite populations mutation frequencies and the onset of Muller's ratchet*. J. Theor. Biol., 137 (1989), 375–395.

- [66] S. Ojosnegros, N. Beerenwinkel, T. Antal, M.A. Nowak, C. Escarmis, E. Domingo. *Competition-colonization dynamics in an RNA virus*. Proc Natl Acad Sci U S A, 107 (2010), 2108–2112.
- [67] K.M. Page, M.A. Nowak. *Unifying evolutionary dynamics*. J. Theor. Biol., 219 (2002), 93–98.
- [68] N. Pariente, A. Airaksinen, E. Domingo. *Mutagenesis versus inhibition in the efficiency of extinction of foot-and-mouth disease virus*. J. Virol., 77 (2003), 7131–7138.
- [69] N. Pariente, S. Sierra, P.R. Lowenstein, E. Domingo. *Efficient virus extinction by combinations of a mutagen and antiviral inhibitors*. J. Virol., 75 (2001), 9723–9730.
- [70] C. Perales, R. Agudo, S.C. Manrubia, E. Domingo. *Influence of mutagenesis and viral load on the sustained low-level replication of an RNA virus*. J Mol Biol, 407 (2011), 60–78.
- [71] C. Perales, R. Agudo, H. Tejero, S.C. Manrubia, E. Domingo. *Potential benefits of sequential inhibitor-mutagen treatments of RNA virus infections*. PLoS Pathog, 5 (2009), e1000658.
- [72] C. Perales, M. Henry, E. Domingo, S. Wain-Hobson, J.P. Vartanian. *Lethal mutagenesis of foot-and-mouth disease virus involves shifts in sequence space*. J Virol, (2011), 12227–12240.
- [73] C. Perales, R. Lorenzo-Redondo, C. Lopez-Galandez, M.A. Martinez, E. Domingo. *Mutant spectra in virus behavior*. Future Virology, 5 (2010), 679–698.
- [74] C. Perales, R. Mateo, M.G. Mateu, E. Domingo. *Insights into RNA virus mutant spectrum and lethal mutagenesis events: replicative interference and complementation by multiple point mutants*. J. Mol. Biol., 369 (2007), 985–1000.
- [75] J.K. Pfeiffer, K. Kirkegaard. *Increased fidelity reduces poliovirus fitness under selective pressure in mice*. PLoS Pathogens, 1 (2005), 102–110.
- [76] M.E. Quinones-Mateu, E. Arts. *Virus fitness: concept, quantification, and application to HIV population dynamics*. Current Topics in Microbiol. and Immunol., 299 (2006), 83–140.
- [77] D.B. Saakian, C.K. Biebricher, C.K. Hu. *Phase diagram for the Eigen quasispecies theory with a truncated fitness landscape*. Physical review, 79 (2009), 041905.
- [78] D.B. Saakian, E. Munoz, C.K. Hu, M.W. Deem. *Quasispecies theory for multiple-peak fitness landscapes*. Physical Review E, 73 (2006), 041913.
- [79] R. Sanjuan, M.R. Nebot, N. Chirico, L.M. Mansky, R. Belshaw. *Viral mutation rates*. J Virol, 84 (2010), 9733–9748.
- [80] D. Shriner, A.G. Rodrigo, D.C. Nickle, J.I. Mullins. *Pervasive genomic recombination of HIV-1 in vivo*. Genetics, 167 (2004), 1573–1583.
- [81] S. Sierra, M. Davila, P.R. Lowenstein, E. Domingo. *Response of foot-and-mouth disease virus to increased mutagenesis. Influence of viral load and fitness in loss of infectivity*. J. Virol., 74 (2000), 8316–8323.
- [82] P. Simmonds. *Recombination in the Evolution of Picornaviruses*. In: The Picornaviruses. E. Ehrenfeld, E. Domingo and R.P. Roos, eds, ASM Press, Washington, D.C. p.p. 229–238, 2010.
- [83] P. Simmonds, S. Midgley. *Recombination in the genesis and evolution of hepatitis B virus genotypes*. J Virol, 79 (2005), 15467–15476.
- [84] H.A. Simon. *The Sciences of the Artificial (3rd edition)*. The MIT Press, Cambridge, Massachusetts, 1996.
- [85] F. Sobrino, M. Davila, J. Ortin, E. Domingo. *Multiple genetic variants arise in the course of replication of foot-and-mouth disease virus in cell culture*. Virology, 128 (1983), 310–318.
- [86] R. Solé, B. Goodwin. *Signs of Life. How Complexity Pervades Biology*. Basic Books, New York, 2000.
- [87] R. Sousa. *Structural and mechanistic relationships between nucleic acid polymerases*. Trends Biochem. Sci., 21 (1996), 186–190.
- [88] D.A. Steinhauer, E. Domingo, J.J. Holland. *Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase*. Gene, 122 (1992), 281–288.
- [89] T.A. Steitz. *DNA polymerases: structural diversity and common mechanisms*. J. Biol. Chem., 274 (1999), 17395–17398.
- [90] J. Swetina, P. Schuster. *Self-replication with errors. A model for polynucleotide replication*. Biophys. Chem., 16 (1982), 329–345.
- [91] J. Sztuba-Solinska, A. Urbanowicz, M. Figlerowicz, J.J. Bujarski. *RNA-RNA recombination in plant virus replication and evolution*. Annual review of phytopathology, 49 (2011), 415–443.
- [92] M.N. Teng, M.B. Oldstone, J.C. de la Torre. *Suppression of lymphocytic choriomeningitis virus-induced growth hormone deficiency syndrome by disease-negative virus variants*. Virology, 223 (1996), 113–119.
- [93] R.C. Valentine, R. Ward, M. Strand. *The replication cycle of RNA bacteriophages*. Adv. Virus Res., 15 (1969), 1–59.
- [94] M. Vignuzzi, R. Andino. in: E. Ehrenfeld, E. Domingo, R.F. Roos (Eds.). *The Picornaviruses*. ASM Press, Washington DC, 2010, pp. 213–228.
- [95] M. Vignuzzi, J.K. Stone, J.J. Arnold, C.E. Cameron, R. Andino. *Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population*. Nature, 439 (2006), 344–348.
- [96] C. Weissmann, M.A. Billeter, H.M. Goodman, J. Hindley, H. Weber. *Structure and function of phage RNA*. Annu Rev Biochem, 42 (1973), 303–328.
- [97] C. Weissmann, T. Tanaguchi, E. Domingo, D. Sabo, R.A. Flavell. *Site-directed mutagenesis as a tool in genetics*, in: J. Schultz, Z. Brada (Eds.). Genetic manipulation as it affects the cancer problem, Academic Press, New York, 1977, pp. 11–36.
- [98] C.O. Wilke, C. Ronnewinkel, T. Martinetz. *Dynamic fitness landscapes in molecular evolution*. Physics Reports, 349 (2001), 395–446.
- [99] S. Wright. *The roles of mutation, inbreeding, crossbreeding, and selection in evolution*. Proc. of the VI International Congress of Genetics, 1 (1932), 356–366.
- [100] M. Yarus. *Life from an RNA world. The ancestor within*. Harvard University Press, Cambridge, Massachusetts and London, England, 2010.

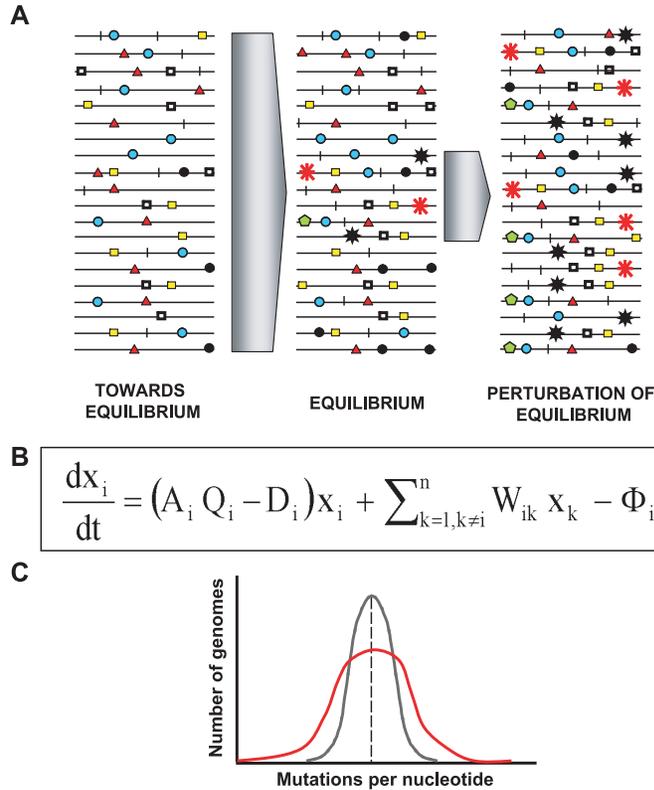


FIGURE 1. Viral quasispecies. A. Mutant spectra are represented by genomes (horizontal lines) with scattered mutations (symbols on the lines). Large population passages in a constant environment, represented here by the broad arrow, tend to drive the population into an equilibrium (middle distribution). Environmental changes or forced replication of only a subset of genomes (small arrow on the right) tend to perturb the equilibrium. Perturbation is diagnosed by an abundance of mutations present in the initial six founder genomes, those drawn preceding the small arrow and that originate the mutant distribution on the right. Periods of near-equilibrium in alternation with alterations of equilibrium are probably the rule during replication of RNA viruses in nature. B. The basic quasispecies equation derived by Eigen and Schuster. It describes the variation of the concentration of mutant i as a function of time (dx_i/dt). A_i and D_i are rate parameters for the replication and degradation of i , respectively. Q_i gives the fraction of correct copies of i produced upon replication of i . W_{ik} describes the synthesis of i as a result of the error-prone replication of template k . ϕ_i indicates the flux of molecules of i away from the environment in which replication takes place. [See [33] for the detailed derivation and implications]. The equation describes a steady-state mutant distribution in equilibrium. Extensions to finite replicon populations in variable environments have been developed, as discussed in the text. C. A graphic representation of two mutant distributions with the same average mutation frequency (given by the discontinuous vertical line) but different amplitude of the mutant spectrum. In the broader (flatter) distribution, extreme components of the mutant spectrum (with a number of mutations higher or lower than the average) occupy a larger portion of the population. The transition from a narrow to a broad distribution may occur as a result of replication of a clonal (originated from a single genome) viral population. Implications of the structure and composition of mutant spectra are discussed in the text.

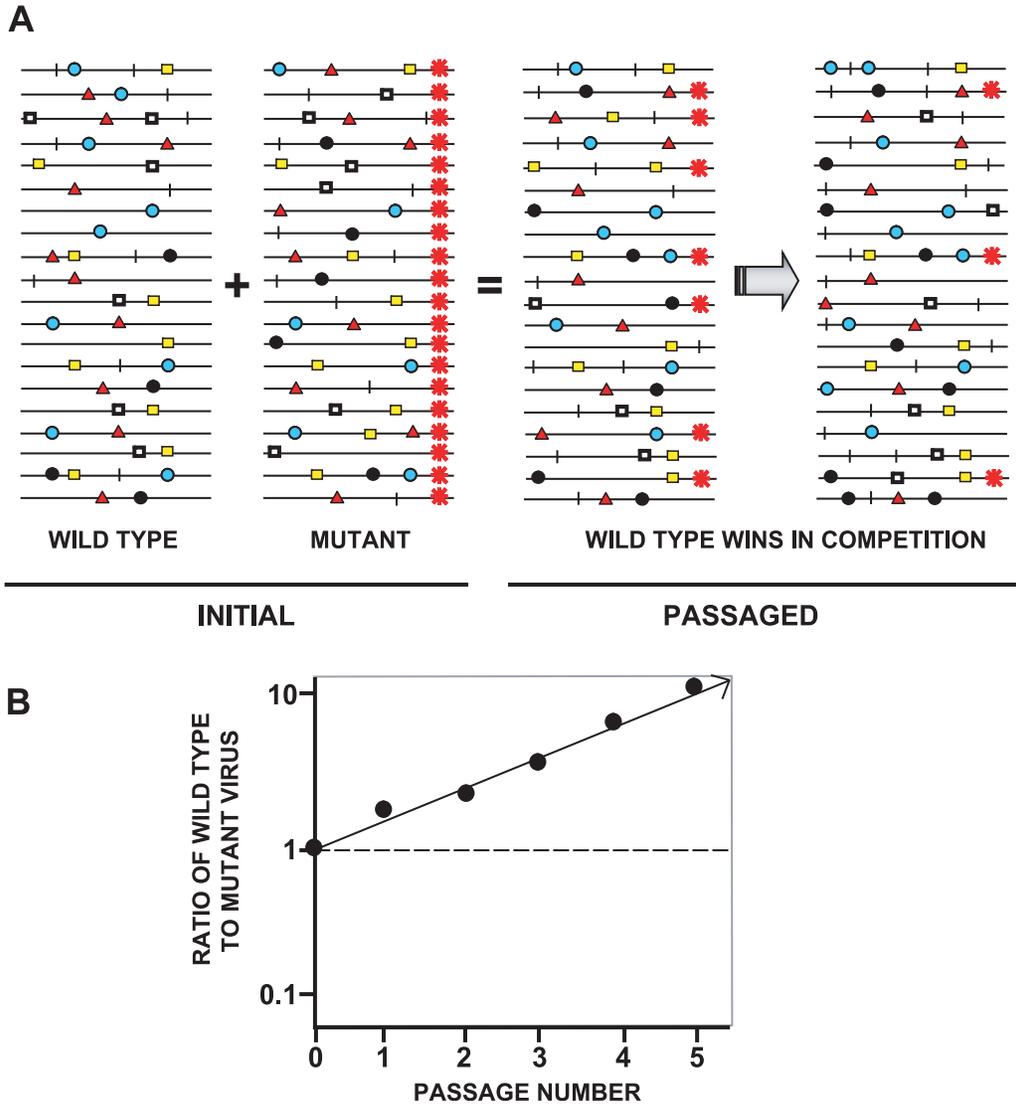


FIGURE 2. Fitness of viral populations. Relative viral fitness is calculated by means of a competition between a wild type (or reference virus) and a mutant to be tested, carried out in a defined physical and biological environment. A. A selective advantage (higher fitness) of the wild type mutant distribution is evidenced by the decrease in the proportion of the mutation that distinguishes the mutant from the wild type (asterisk). B. Plot used to determine the fitness vector. In this logarithmic plot of the ratio of wild type to mutant virus (ordinate) versus passage number (abscissa with linear scale) the antilogarithm of the slope of the fitness vector (regression line of the experimental points, circles) is the fitness value of the wild type relative to the mutant. See text for references.

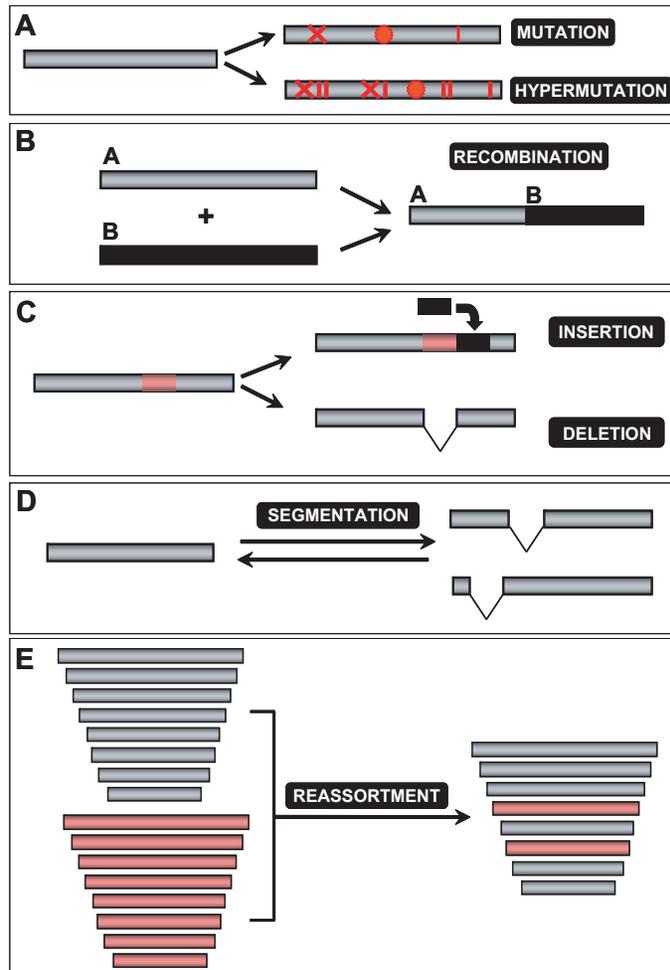


FIGURE 3. Schematic representation of common types of genetic variation of viral genomes. A. Mutation due to limited fidelity of viral polymerases; symbols on the genome indicate different mutation types: four transition ($A^- \rightarrow G$, $G^- \rightarrow A$, $U^- \rightarrow C$, $C^- \rightarrow U$) and eight transversions ($A^- \rightarrow U$, $A^- \rightarrow C$, $G^- \rightarrow U$, $G^- \rightarrow C$, $U^- \rightarrow A$, $U^- \rightarrow G$, $C^- \rightarrow A$, $C^- \rightarrow G$). Hypermuted genomes (often with a biased mutation type) are generated by the action of cellular editing activities such as APOBEC and ADAR, which are diverted from their physiological role into acting as an innate immune defense against some viruses. B. Recombination is the formation of a mosaic genome from two parental genomes. There are several types of recombination: homologous, non-homologous, replicative and non-replicative. C. Insertions or deletions (abbreviated “in-dels”) can be viewed as the result of a type of intragenome or intergenome recombination that may involve non-viral genomic sequences. They are frequent in the generation of defective-interfering particles, abundant in negative-strand RNA viruses. D. Internal deletions can give rise to a segmented form of a genome that replicates and kills cells by complementation. E. Segmented genomes often undergo segment reassortment, as in the antigenic shift of influenza viruses. None of the types of genetic variation are mutually exclusive: reassortment, recombination, mutation and hypermutation can act simultaneously on a segmented viral genome although such divergent genomes must be rare and unfit. Schemes based on [17, 22].

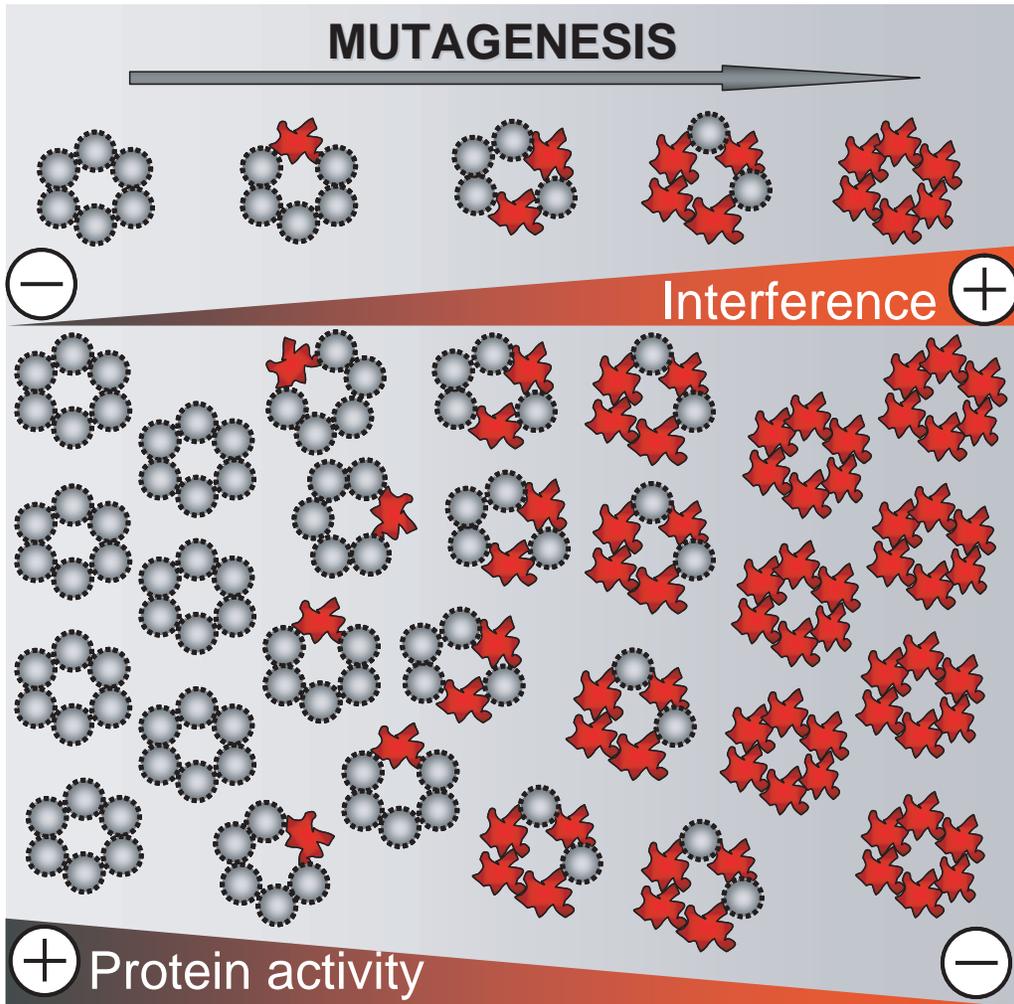


FIGURE 4. A schematic representation of the interference exerted by an altered viral protein that functions as a hexameric complex. On the left hexamers made of smooth grey subunits are functional and protein activity is high. As additional mutations are introduced in the genome, defective forms of the protein (represented with an irregular surface) increase in frequency. They interfere by forming either inactive hexamers (made of rough subunits) or partially active proteins (hybrid hexamers). This model is based on experimental and *in silico* results described in the text.

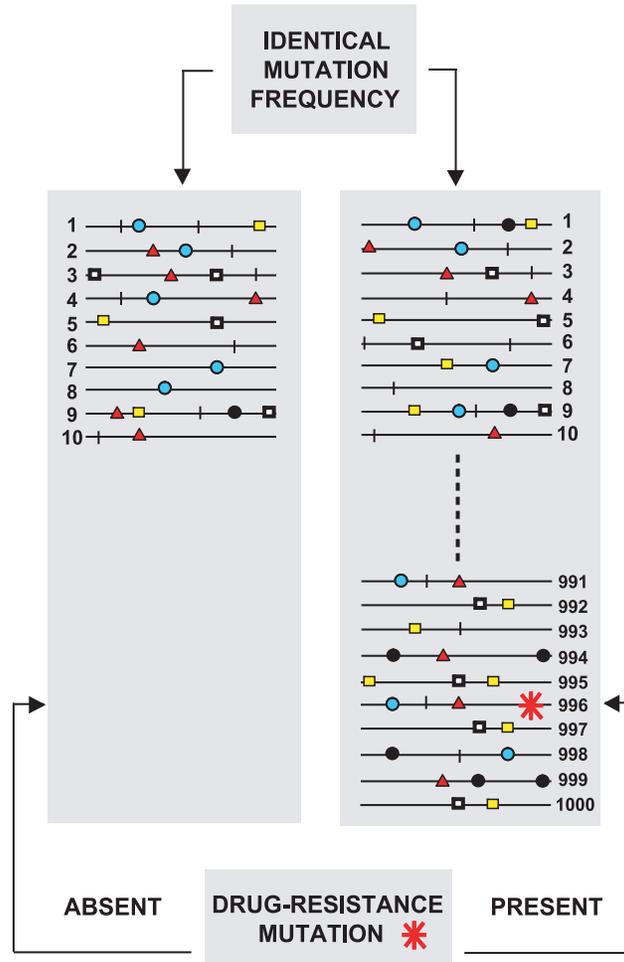


FIGURE 5. **The weight of viral load on the adaptive potential of a viral genome population.** Mutant distributions are depicted using the same symbols as in Figures 1 and 2. Both of them (10 genomes on the left and 1000 genomes on the right) have been sampled from the same quasispecies and, therefore, they have the same mutation frequency. A specific drug resistance mutation (at the genomic site indicated by an asterisk) is present in the large distribution but not in the small one. While for a replicative unit mutation frequency is an “intrinsic” property of the distribution (independent of the sample size taken), the presence of specific variants depends on the sample size (“extrinsic” property). Population size and genetic heterogeneity are key parameters in quasispecies behavior. See text for further comments and relevant references.