Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physics of Life Reviews 3 (2006) 65–92
PHYSICS of LIFE reviews
www.elsevier.com/locate/plrev

Viral evolution

Susanna C. Manrubia *, Ester Lázaro

Centro de Astrobiología, Ctra. de Ajalvir km. 4, 28850 Torrejón de Ardoz, Madrid, Spain

Accepted 17 November 2005

Available online 6 January 2006

Abstract

In the last two decades, viruses have been used as model systems to study evolution in short periods of time. Due to their characteristics, virus adapt rapidly to changing conditions, thus allowing the quantification of several evolutionary features under controlled laboratory conditions. Here we review the basic biology of viruses and describe in detail a number of experiments performed with RNA viruses. Particular emphasis is devoted to the interpretation of the experiments and to the involved phenomenology. This analysis sometimes represents the basis to formulate simple evolutionary models that aim at describing the observed dynamics. In other cases, theoretical results have prompted the realization of related experiments, as we discuss. Concepts as fitness loss and fitness recovery, the error threshold, increased mutagenesis, viral sex, or viral competition and interference, are discussed in an empirical framework and from the associated theoretical point of view.

PACS: 87.23.Kg; 87.14.Gg; 87.10.+e

Keywords: RNA virus; Molecular evolution; Quasispecies theory; In vitro experiments; Mathematical model; Evolutionary strategy

High error rates, large populations, and short generation times are the three features that have made of viruses a suitable system to witness evolution. The prospects opened by the direct observation of evolution at the population level, and the possibility of tracing adaptation to its molecular roots are difficult to overstate. In the last two decades, the number of experiments involving viruses and addressing fundamental questions on the mechanisms of mutation and adaptation has grown enormously. Still, there is a long way to go before we fully apprehend the laws governing molecular evolution. Often, empirical results defeat theoretical expectations, and our assumptions on the dynamics underlying evolution have to be revised over and over.

This review is aimed at researchers interested in the biological, phenomenological, and conceptual issues involved in viral evolution. The work is divided in three parts. We begin with a relatively detailed introduction to the biology of viruses, where we briefly review their discovery, the general structure of a virus, and its reproductive cycle inside the cell. The first part continues with a description of adaptive mechanisms acting at the level of the viral genome and of properties of quasispecies. Particular emphasis is devoted to the concept of fitness, a notion that cannot be defined without taking into account the environment where evolution proceeds. Table 1 contains glossary of the most common terms used in viral evolution.

* Corresponding author.
E-mail address: cuevasms@inta.es (S.C. Manrubia).

1571-0645/$ – see front matter
doi:10.1016/j.plrev.2005.11.002

Table 1
Glossary of terms and acronyms frequently used in virology

Virus	Infectious agent composed by a genomic nucleic acid covered by several protective layers. The indispensable condition for its multiplication is the infection of a host cell that provides the necessary resources for genome replication and assembly of daughter particles
Viral genome	Nucleotide sequence that contains the virus genetic information
Virion	Mature virus particle. This term is usually used for the extracellular phase of the viral cycle
Viral clone	Viral population obtained upon replication of a single parental virus
Bacteriophage	Virus whose natural host is a bacteria
Fitness	A parameter that quantifies the adaptation of an organism to a given environment. It is a relative value and might involve a number of phenotypic traits
Consensus sequence	Sequence resulting from taking, for each position, the most frequent residue (nucleotide or amino acid) found in the set of genomes of the population. The consensus sequence may not exist physically in the mutant spectrum. This is the sequence yielded by usual sequencing protocols
Defective viral particle	Virus particle whose genome lacks some essential portion necessary to complete an infective cycle. Some defective particles can give rise to productive infections when they are able to use the resources produced for other viruses that replicate inside the same cell. This is called trans complementation
Lytic infection	Infection produced by a virus that replicates through a cycle in which progeny virions destroy the infected cell
Persistent infection	Infection where the host cell remains alive and produces progeny virions often at a slow rate, but for a long time
Mutation rate	Frequency of occurrence of a mutational event during genome replication
Mutational load	Total number of mutations accumulated in a population
Viral load	Number of infectious particles in a viral population
Epistasis	Interactions between different mutations
Multiplicity of infection (MOI)	Average number of infectious virus per susceptible cell
RT	Reverse transcriptase. Protein copying RNA into DNA. It is used by retroviruses, such as HIV
PCR	Polymerase chain reaction
PFU or pfu	Plaque forming units. It quantifies the infectivity of a virus
HIV	Human Immunodeficiency virus
VSV	Vesicular Stomatitis Virus
FMDV	Food-and-mouth disease virus
LCMV	Lymphocytic choriomeningitis virus
SARS	Severe acute respiratory syndrome

The second part of the review deals with empirical observations of viral evolution. It starts with some examples of viral evolution in nature, and discusses the difficulties of interpretation inherent to evolutionary patterns in that case. A section detailing laboratory protocols precedes the detailed description of several different experiments with RNA viruses, mostly. Though theoretical approaches to evolution are especially tackled in the third part of the review, it is impossible to fully separate the interpretation of certain experiments from the conceptual elaboration of the results and, in some cases, from the mathematical models they have inspired. In this sense, the second and the third part are deeply intermingled.

The literature on viral evolution, and in particular on theoretical models dealing with viral evolution, is vast. We have decided to restrict this revision mostly to RNA viruses and to models that have been either directly inspired by the experiments or that are conceptually relevant in our understanding, such that they inspired experiments in their turn. We regret that this decision excludes many interesting publications on general quasispecies theory, for instance.

We touch only briefly on epidemic processes in nature, but leave aside the wealth of publications on related theoretical aspects. Our main effort has been directed towards presenting the reader with an as complete as possible revision of the empirical observations on viral evolution and its current interpretation.

1. Introduction to viruses

By the end of the 19th century, it became clear that the spontaneous generation of organisms did not occur [77]. The efforts of many scientists, among them Louis Pasteur, Robert Koch and Joseph Lister, were necessary to eradicate a belief that had been the subject of strong debate during decades. The end of spontaneous generation gave birth to the notion that life only comes from life. One of its consequences was the enunciation of Koch's postulates for defining whether an organism was the causative agent of a disease.[1] Koch's postulates became a paradigm in microbiology and brought about the design of sterilization methods that allowed to cultivate pure organisms. One of those methods consisted in the use of filters that most known bacteria could not go through. In this context, the discovery by Adolf Mayer and Dimitri Ivanofsky of a new infectious agent, the causative of tobacco mosaic disease, which was able to go through filters retaining infectivity had to fight against new and old dogmas. The agent could not be isolated in pure culture, contradicting Koch's second postulate. At first, it was supposed that either the filters were defective or that this "filterable agent" could be a toxin. However, Martinus Beijerinck soon made the crucial discovery on the ability of this agent to reproduce itself, though only inside living cells. For the first time, the existence of replicating entities smaller than bacteria, not observable with a light microscope, became clear. These entities were usually referred as *filterable agents*, and it was only after the first decades of the 20th century that the term virus started to be used with the same meaning that it has today. The discovery of viruses constitutes a good example of how knowledge in biology often defeats well established theories and requires their reformulation in order to include the new discoveries.

The first virus from animals, foot-and-mouth disease virus (FMDV) [79], was described and isolated shortly after. The existence of viruses able to infect bacteria, the so-called bacteriophages, was discovered in 1901. The use of bacteriophages was crucial to develop methods and techniques that are currently used in virology and in the analysis of other important life processes, for instance the transmission of genetic information. Studies with phages such as the T series, ΦX174, Qβ, or λ, have also permitted to investigate in a very detailed way the processes going on inside virus-infected cells, and to establish a sequence of events that, with some differences for each particular virus, must be successfully completed to produce an infectious progeny. The isolation of different variants of the same virus was another crucial discovery: viruses were able to mutate and adapt to changes in the environment obeying the same laws of natural selection that Darwin had enunciated for higher organisms [88].

At that point, the ingredients for inaugurating a new debate, namely, whether viruses are living entities, were all available. On the one hand viruses possess a genetic material encoding a (usually low) number of proteins and able to replicate. On the other hand, viruses lack the metabolism required to perform the replication of their genomes and the translation of the encoded information into proteins. They are strictly dependent on the enzymatic activities and the energy sources supplied by the infected cells: they are obligate intracellular parasites. The debate is still open, and relies heavily on the definition of "life" used.

1.1. The viral structure

A virus is formed by one to several nucleic acid molecules protected by a closed shell, the capside, formed by proteins. In some viruses, the capside is surrounded by a lipid bilayer furnished by the host cell. The nucleic acid is the genetic material of the virus, the genotype, which encodes the viral proteins. The genotype determines the features of the new progeny viruses that will be produced inside the infected cell. These features constitute the phenotype, and usually result in adaptive differences. At odds with all cellular organisms, which always use DNA to store the genetic

[1] The four postulates formulated by Koch in 1884 are: 1. The organism must be found in all animals suffering from the disease, but not in healthy animals; 2. The organism must be isolated from a diseased animal and grown in pure culture; 3. The cultured organism should cause disease when reintroduced into a healthy animal; 4. The organism must be reisolated from the experimentally infected animal.

Table 2
Different types of viral genomes

Genome type	Viral families	Examples
dsDNA	*Adenoviridae, Herpesviridae, Papillomaviridae, Polyomaviridae, Poxviridae*	Herpes, phage T4, phage λ, mimivirus, chicken pox, varicella virus
ssDNA	*Parvoviridae*	Phage ΦX174, canine parvovirus, chicken anemia virus, chlamydia phage 1
RNA and DNA reverse-transcribing viruses	*Hepadnaviridae, Retroviridae*	HIV, hepatitis B virus, cauliflower mosaic virus, avian leukosis virus, human spumavirus
dsRNA	*Reoviridae*	Phage Φ6, infectious pancreatic necrosis virus, simian rotavirus SA11, drosophila X virus
Negative-sense ssRNA viruses	*Arenaviridae, Bornaviridae, Bunyaviridae, Filoviridae, Orthomyxoviridae, Paramyxoviridae, Rhabdoviridae*	Measles, influenza, rabies, Ebola virus, LCMV, VSV
Positive-sense ssRNA viruses	*Arteriviridae, Astroviridae, Caliciviridae, Coronaviridae, Flaviviridae,* Hepatitis E-like viruses, *Picornaviridae, Togaviridae*	Poliovirus, SARS coronavirus, FMDV, tobacco mosaic virus, rhinovirus (causing common cold), hepatitis A, rubella, dengue, and yellow fever virus, Qβ phage

Source: Fields VIROLOGY, 4th edition.

information, viral genomes can be composed of DNA or of RNA. This establishes a basic distinction between DNA viruses and RNA viruses. In both categories of viruses it is possible to find examples in which the genetic material is present as single-stranded molecules (ss) or as double-stranded molecules (ds) (Table 2).

To transmit the genetic information to the next generations the nucleic acid has to be replicated, a process that involves cellular and viral enzymes. Usually, DNA viruses replicate their genomes with similar enzymatic activities used by the host cell. In contrast, RNA viruses are the only entities in nature where RNA is used as repository of genetic information. This implies the design of replication mechanisms exclusively employed by this type of viruses. RNA virus replication is a very error prone process [34], a feature with relevant consequences in the evolution of RNA viruses and in the amount of information that can be stored in their genomes. Measurements of the chain lengths and of the replication error rate of RNA viruses show that the genome length is close to the maximum that can be faithfully maintained. A consequence is that RNA viruses possess highly compacted genomes, with regions sometimes involved in several functions (as an example, coding regions can act also as regulatory signals).

The proteins encoded by the viral genomes can be grouped into several basic categories that include structural proteins (later constituting the protective capside or the external spikes, e.g.), enzymatic activities (necessary for the replication of the virus genome), regulatory proteins (required for the correct assembly of the progeny viruses), and other proteins essential to perform processes that are virus dependent, such as the release of virions from the infected cell. In addition to the composition of the nucleic acid, another basic characteristic to classify viruses is the presence of the external membrane, which distinguishes enveloped viruses from non-enveloped ones. Most enveloped viruses acquire the envelope by budding through a membrane of the host cell, a step that facilitates the extrusion of viruses from the infected cell.

1.1.1. Reproductive cycle of viruses

Though there are large differences in the reproductive cycles of different virus, the basic processes that any virus must complete are as follows.

1. *Binding to specific cell surface receptors*. These receptors are major determinants of cell tropism (the cellular types and organisms that can be infected by a given virus) [4]. In addition to primary binding receptors, some viruses also interact with secondary receptors (named coreceptors) that collaborate in the penetration of the virus inside the cell. A comparison of known viral receptors and coreceptors reveals that, occasionally, members of the same virus family have evolved to use a variety of different cell surface proteins for viral entry. It is also possible that very distinct viruses use the same receptor.

2. *Entry of the virus into the cell*. Viruses can enter the cell in two main forms: as naked nucleic acid molecules that carry all the information necessary to originate new viruses or as complete particles. The entry of enveloped viruses requires the fusion of viral and cellular membranes by a process driven by viral glycoproteins located on the virus surface. Non-enveloped virus entry into cells involves the disruption of the cellular membrane, in some cases through the formation of pores.
3. *Viral uncoating*. This process consists in the disorganization of the protective protein layers to release the viral nucleic acid inside the cellular cytoplasm. Sometimes these events are coupled with penetration through the cellular membrane, whereas other viruses experience the disassembly steps inside the cytoplasm.
4. *Replication and expression of genetic information*. Briefly, replication of RNA viruses occurs through two basic mechanisms that distinguishes riboviruses from retroviruses [3]:
   - RNA-dependent RNA synthesis. This process takes place in riboviruses and it is catalyzed by the enzyme RNA replicase.
   - RNA-dependent DNA synthesis. This is a process known as reverse transcription that is catalyzed by the enzyme reverse transcriptase of retroviruses.

   Both enzymes, RNA replicases and reverse transcriptases, have to be encoded by the viral genomes and expressed early during the infection cycle. In some cases, these enzymes are copackaged with the viral nucleic acid during the assembly of the viral particles.

   The intermediary molecule in the flux of genetic information that leads to the synthesis of specific proteins is the messenger RNA (mRNA), the RNA molecule accepted by the cellular ribosomes. This establishes a basic distinction between RNA viruses with negative or positive polarity. Positive polarity viruses can use their genomic RNAs as mRNAs that can be directly translated. In negative stranded viruses the genomic RNA has to be copied to produce subgenomic mRNA and full-length antigenomic RNA. The former can be translated, and the latter can be replicated again to obtain new negative-sense RNA genomes. Retroviruses represent an exception to this scheme. They possess positive polarity genomes that, instead of being translated directly, have to be copied to DNA before the replication cycle is completed. The newly synthesized DNA is integrated into the cellular DNA and transcribed by a cellular RNA polymerase to produce new RNA molecules. The latter can be used as mRNA or packaged inside the new virions.
5. *Virus assembly*. This is the process by which a number of structural and non-structural proteins are assembled together with the genomic RNA to constitute a new virus particle.
6. *Virus release*. This a crucial step to propagate the infection cell-to-cell or organism-to-organism. Different viruses employ different strategies to leave the cell. Enveloped viruses (those enclosed by a lipid bilayer) are usually released by budding from the plasma membrane. Non-enveloped viruses generally escape via lysis of the infected cell.

1.2. Mutation, selection, and adaptation

Evolution through natural selection necessitates the generation of a certain genetic diversity on which it can act. This diversity is eventually responsible for the existence of differences in the reproductive success of individuals, in their fitness. Most genetic variability is a direct consequence of the occurrence of errors during the copy of viral genomes. Nucleic acid replication is a complex process where a complementary sequence is created through the sequential addition of nucleotides. Replication is not fully faithful and, occasionally, an incorrect nucleotide is added to the synthesizing chain. If the error is not corrected, it can be propagated to the next generations and results in a mutation. Apparently, there might be a contradiction between the deleterious effect of most mutations and the need of diversity to permit adaptation to changing conditions. Comparison of the spontaneous mutation rates of replication in different organisms suggests that evolution has selected optimal mutation rate values as a function of the environmental variability [33,38].

DNA genomes are replicated by DNA polymerases, enzymes endowed with several corrector activities that ensure a high copying accuracy. The accuracy of DNA polymerases can be however modified if the environment requires it. For example, stressing conditions can promote the selection of bacterial variants that lack some of the corrector activities of the polymerase [123]. These variants behave as hypermutator strains, thus generating a broader genetic diversity necessary to promote the appearance of individuals able to reproduce efficiently under the stressing conditions. The immunoglobuline genes of vertebrates constitute another example of DNA replicating with an error rate

higher than that of the remaining genes in the genome. The extra variability is essential to produce a very large number of antibodies able to neutralize a large diversity of antigens.

Viruses are replicators that employ the general mechanisms of mutation and selection to reproduce efficiently in frequently changing environments. There are important differences in the evolutionary capacity of DNA and RNA viruses. DNA viruses can replicate their genomes using high fidelity DNA polymerases. In contrast to this, and because in the cellular world RNA replication does not take place (RNA is always synthesized upon transcription of DNA), RNA viruses need enzymatic activities that are not usually present in the cell. As previously mentioned, these enzymes are the RNA replicases and the reverse transcriptases, which are encoded by the viral genomes. Both enzymes lack corrector activities, resulting in an error rate of replication several orders of magnitude higher than that affecting DNA replication, permitting eventually the generation of a huge variety of mutant genomes [29,34].

The particularities of RNA virus replication and their consequences in adaptation have a great relevance in the emergence of new diseases in humans and in the possibility of eradicating virus infections caused by RNA viruses. Due to the complex behavior exhibited by RNA viruses [91], the rest of this review will be devoted to their evolution.

1.2.1. Evolutionary mechanisms in viral genomes

The average mutation rate of RNA viruses ranks in the order of 10^{-4} to 10^{-5} misincorporations per nucleotide and per round of copying [5]. This means that each daughter genome will contain on average one or two mutations when compared to the parental sequence. Mutation rate must be distinguished from the mutation frequency, which is the fraction of mutations in a genome population. This last parameter is influenced by the replicative ability of each mutant genome in competition with the rest of genomes in the population. Positive selection favours the presence of mutations with an adaptive value whereas negative selection purges or maintains in low amount genomes carrying deleterious mutations. The frequency of a given mutation in the population can also be influenced by genetic drift, which results from changes in the population size that may lead to the stochastic fixation of mutations irrespectively of their selective value. The main result of genetic drift is a random sampling of the genomes rendering a new composition of the population.

In addition to point mutations, other processes that contribute to modulate the genetic diversity in RNA viruses are recombination and genome segment reassortment (Fig. 1). Recombination consists in the generation of a new mutant genome by linking two or more fragments belonging to different parental molecules. Recombination may be a powerful mechanism to accelerate evolution by the acquisition of groups of mutations present in different genomes. However, there are not enough studies performed to date to evaluate the role played by recombination in RNA virus evolution [12,102,105].

Genome segment reassortment takes place in viruses with segmented genomes, such as influenza A, and consists of the encapsidation in the same viral particle of genomic segments belonging to two or more different parental viruses [37]. The resultant genetic combinations can confer new properties that strongly increase virus diversification. In the case of influenza A, this type of genetic exchange is known as genetic shift and has been responsible for the emergence in humans of viruses carrying genes from avian or pig influenza strains. Occasionally, the new genes prevent recognition of the influenza virus by the immune system of the host, which can result in increased virulence. If these reassortant viruses can be transmitted person-to-person, an influenza pandemic can probably result.

1.2.2. Viruses and molecular quasispecies

All the mechanisms generating genetic diversity in large populations of fast replicating individuals, as RNA viruses, lead to a highly heterogeneous population structure. When the environment experiences a change, the most successful variants in the new conditions increase their representation in the population, thus changing its composition. Once the processes of mutation, competition and selection have acted for a long enough time in a constant environment, a stationary state where each mutant is represented in a fixed amount, is reached. This stationary state defines a quasispecies and represents the counterpart in viruses of the theoretical quasispecies defined in the context of prebiotic evolution of populations of RNA replicators subjected to mutation and selection [41,42] (see also Section 3.1).

The isolation of variants of tobacco mosaic virus with morphology different from the standard type was reported as early as in 1926 [88]. Few years later, Jensen [69] confirmed those observations and further detected the presence of revertant mutants. An estimation of the degree of heterogeneity in RNA virus populations could only be determined once nucleotide analysis techniques were available for RNA. Esteban Domingo and colleagues, working with clones of the bacteriophage Qβ, estimated the diversity of fragments obtained after digestion with RNAase T1 (an enzyme

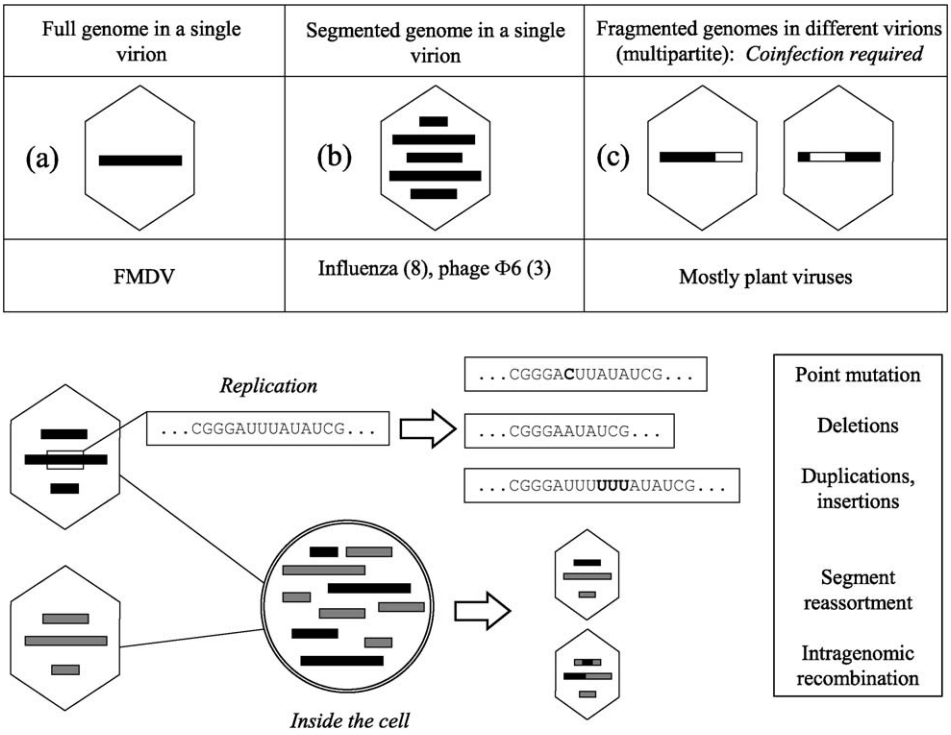

Fig. 1. Different forms of viral genomes in nature and their evolutionary strategies. (a) A complete genome is encapsidated in a single virion. (b) Other viruses present segmented genomes, with a fixed number of fragments forming the viable virion (shown in brackets). (c) In certain cases, uncomplete genomes are encapsidated in different viral particles, such that coinfection of the same cell by both types is needed to cause infection. In all cases, point mutations, deletions and nucleotide insertions in the genome (be it in single molecules, segmented or multipartite), as well as recombination inside the cellular interior are used as evolutionary strategies. In case (b) reassortment of segments is possible. In both cases (b) and (c), a form of viral sex may hold inside the cell. Some examples of natural viruses with each type of genome are given. In Section 2.4.2, a transition from type (a) to type (c) under an environmental change is described.

that cut the RNA at specific sequences of nucleotides). They found that about 15% of the clones derived from a multiply passaged Qβ population showed patterns which deviated from those obtained with the bulk RNA extracted from the total virus population [32]. Correction of this value by the length of the sequence analyzed by this technique leads to an estimate of one or two nucleotide differences in each phage genome. As Esteban Domingo and colleagues posed it: *The genome of Qβ phage cannot be described as a defined unique structure, but rather as a weighted average of a large number of different individual sequences* [32]. These early observations have been confirmed by sequencing techniques and have been extended to all RNA viruses studied.

The high level of genetic diversity that is present in quasispecies leads to a more diverse and sometimes unexpected behaviour than that observed in less variable populations. A viral quasispecies behaves as a dynamic system in continuous turnover where the processes of generation of new mutants, selection of the best adapted and elimination (or maintaining at low amounts) of the less fit genomes is always taking place. The determination of the sequence of an RNA virus population yields a sequence that corresponds to the most represented nucleotide at each genomic position. This is the consensus sequence of the population, and does not represent the high genomic diversity of a quasispecies.

An interesting hypothesis supported by several experimental observations is the existence of different levels of organization inside the quasispecies, in the sense that different genomes are present in different amounts, and simultaneously linked to each other. Genomes that are only present in low amounts can be determinant in the evolution followed by the population in response to an environmental change. Recently, a type of memory in viral quasispecies has been described. This memory refers to genomes that were very abundant in the population at previous evolutionary stages, and that are present in very low amounts at present [31,109]. These relic genomes might be very durable and constitute a reservoir of evolutionary possibilities, previously explored by the virus, which may be used again if an environmental change requires it.

The rate of virus evolution in nature can be defined as the rate of fixation of mutations in the consensus sequence. However, the quasispecies might occasionally experience deep reorganizations that are not reflected in the consensus sequence. Variants at memory levels are an example, since those genomes carry mutations that, in spite of never being detected in the consensus sequence, can have a profound influence in the adaptation of the population. When the environmental conditions are relatively constant the virus population can be maintained in evolutionary stasis, with the result that the consensus sequence keeps almost invariant through time. This situation usually occurs in viruses infecting the same host during very long time periods, as influenza in birds or hantavirus in rodents. Evolutionary stasis does not mean, at short evolutionary time scales, that the mutation rates decrease: It represents instead an equilibrium between mutation and selection that yields a consensus sequence mostly without changes. A virus that is kept invariant for a long time can experience an acceleration in the fixation of mutations when the environment is altered, as it occurs during the infection of a new host. A recently studied example is the coronavirus causing the SARS epidemic [107]. At the beginning of the spread of the virus in the human population, it experienced a high rate of fixation of mutations that probably were essential for the adaptation of the virus to the new host.

1.3. Fitness, trade-offs, and environment

Viral fitness is defined as the relative ability of a virus population to produce infectious progeny under a set of defined environmental conditions [29]. An absolute ranking of fitness values does not exist, and these are usually determined through comparison with a reference virus that is assigned a fitness value of 1 (see Section 2.2).

The concept of fitness in RNA viruses has been quite often simplified in excess, probably due to the direct translation of that concept from molecular evolution theories, where the replicative ability is the only property defining the phenotype. In spite of being relatively simple entities, viruses are much more complex than plain replicators. The reproductive success of a virus depends not only on its ability to replicate its genome: there are many other processes that have to be completed to eventually generate a new infectious virus progeny. These processes, detailed in Section 1.1.1, include the interactions derived from the production in the same cell of a pool of viral proteins and nucleic acids carrying different mutations. This complex mixture of molecules establishes a network of interactions where different genomes play different roles. As a side result of complementation among genomes, it is common to find parasites (genomes that use for their replication resources supplied by other viruses) or altruistic behavior (genomes that produce proteins that are used by other viruses). We can think of a viral population as a kind of entangled ecosystem composed by molecules with different capacities, in continuous interaction, and with a common origin. The emerging behavior of such ensembles is not easy to predict [125].

For example, and contrary to naïve expectation, the genome with the highest genomic replicative ability is not necessarily the most represented in the population. Suppose that a genome replicates fast but lacks the signals necessary to be packaged correctly in the capside. This genome will be very poorly represented in the final population and may even disappear in the long term. This reveals that fitness is a many-fold feature that includes not only the replicative ability of the genome, but also several other processes that must be successfully completed if a viable infectious progeny is to be produced. Optimisation of fitness requires a compromise between the different features involved. For instance, a fast replicator might be at advantage when competing inside the cell with other genomes, but suppressing the ability to correctly encode packing signals brings about a clear disadvantage to propagate. This establishes trade-offs between different features, and the problem of fitness optimisation in viruses turns out to have a number of constrains whose relative relevance depends on the present environment. In order to ensure the long-term survival of a population, all selection pressures acting on the relevant phenotypic traits have to be independently maintained. Otherwise, some functions can be lost and extinction might ensue.

As a consequence, virus fitness is not a constant absolute value: It depends on the precise environment, and external perturbations can completely modify the quantitative fitness value of an individual or of a population. As has been described in the previous section, virus populations are dynamical, heterogeneous ensembles of genomes. The repertoire of genetic variation enables that variants with higher reproductive success in the new environment, if present, increase in number. The selected genomes can become dominant in the population, a situation advantageous for the virus, but that strongly difficults its eradication. The emergence of drug-resistant mutants during the treatment of many diseases produced by RNA viruses [114] is a well-known example. Initially, the drug prevents the replication of most genomes, but the rapid emergence of variants carrying mutations that resist its action causes the failure of many treatments.

Another relevant factor in evolution is population size. In virology, this parameter refers to the number of individuals that will form a new population. When this number is very low, the mutations present in the founder genomes are transmitted to most of the progeny viruses, resulting in an enhanced fixation of mutations in the consensus sequence. Reductions in the population size imply that the processes of competition and selection of the best adapted phenotypes take place among a reduced number of genomes. As most mutations have deleterious effects on fitness, the result of repeated reductions in the population size is usually a decrease in the fitness of the population. In contrast to this, large population passages of a virus population permit competition among a large number of genomes, increasing the probability of fixation of advantageous mutations, which results in fitness increases.

1.3.1. The relationship virus–host

The molecular mechanisms by which a virus can give rise to infectious progeny inside a particular type of cell or host constitutes an amazing example of adaptation, where the virus uses the cellular metabolism for its own benefit. But this use of the cellular resources frequently implies an injury to the cell, often causes disease, and occasionally the death of the infected organism. This has promoted the coevolution of virus and cells, and the latter have developed defense mechanisms that find its more refined expression in the immune system of vertebrates.

In this arms race with their hosts, viruses have adopted different strategies that permit their survival in the long run. Some viruses, such as measles or smallpox, have specialized to multiply inside a single host, whereas others, such as influenza, can infect an ample repertoire of organisms. Other important distinction separates viruses that are maintained in a single host from others that alternatively infect different species, thus establishing complex cycles in which one of the hosts behaves as a vector (usually an insect) where the virus multiplies. The vector transmits the virus to a different host that can represent a dead-end for the virus. Examples of viruses with a cycle involving two or more hosts are arboviruses, arthropod-borne viruses that usually thrive by continuously cycling between an insect host and a vertebrate host.

To perpetuate in the host population, some viruses have specialized in producing short-duration acute infections. At times, the virus can cause serious damage or even the death of the host, but the survival of the virus is guaranteed if the susceptible host population is renewed frequently enough. The time scales involved, and thus the success of this strategy, depend on the duration of the immune response against the virus. An example of this category of viruses is influenza, which through a continuous change in their antigenic properties (antigenic drift) evades the immune system of the host [37]. Usually, this escape mechanism is limited by the fact that, in most viruses, their antigenic determinants are the same proteins required for the penetration of the virus inside the cell [4]. The opposite strategy is that of viruses that have attenuated its virulence up to a limit compatible with persistence in their hosts, without provoking serious disease symptoms. Many of these viruses can stay unnoticed, being only exposed when they infect a new host lacking effective defense mechanisms to defeat the new infectious agent. Acute infections and asymptomatic infections are two extremes of a broad spectrum of relationships between viruses and their hosts.

In general, whenever a virus crosses species boundaries and infects a new host, it is not able to multiply efficiently in the new cellular type. In this case the infection represents a dead end, and the transmission of the virus in the new species halts. However, due to its enormous adaptive capacity, virus sporadically acquire the ability to transmit between organisms, with the result that a new disease emerges.

Transference of viruses is usually favoured by ecological disturbances that increase the contact between species carrying different types of viruses. Sometimes, the transfer is also facilitated by changes in the properties of the virus that permit penetration and replication in a new cellular type. Most recent disease emergences in humans include severe acute respiratory syndrome (SARS, caused by a type of coronavirus without noticed infection in humans before 2003 [107]), acquired immunodeficiency syndrome (AIDS, caused by the retrovirus HIV-1), which is believed to have its origin in a virus whose natural host is a species of chimpanzee [56], and the influenza virus H5N1 (an avian virus that can also infect humans). This last virus is considered by the World Health Organization the most probable candidate to cause an influenza pandemic, if it acquires the necessary abilities to be transmitted person-to-person [130].

2. Empirical observations of viral evolution

The mechanisms underlying viral replication, change, and eventual adaptation are reflected in complex dynamical patterns at the population level. While viral evolution in nature is extremely difficult to interpret, current laboratory

techniques permit to carry out in vitro experiments under well controlled conditions. This opens the possibility of quantifying viral evolution, and represents a first step towards formalization of evolutionary mechanisms. Nowadays, information coming from mutations in the sequence, subsequent modifications in the molecular structure, and variations in the phenotype of individuals, together with dynamical patterns at the population level, can be integrated to yield a detailed picture of the main mechanisms involved in evolutionary change.

2.1. Dynamics of virus evolution in nature

Interactions between viruses and hosts either at the individual or at the population level have a profound influence in the evolution of RNA viruses. Part of this evolution is reflected in the topology of the corresponding phylogenetic trees. The type of infection caused by the virus (acute or persistent), the duration of the immunity elicited by the pathogen, the rate of the replenishment of susceptible hosts, the capacity of the virus to acquire immune-escape mutations, or the transmission mode, are some of the factors that affect the evolutionary dynamics of RNA viruses. Recent developments in molecular biology and bioinformatics have permitted to reconstruct the phylogenies of several RNA viruses along a number of years. The comparison of the phylogenetic trees obtained reveals important differences that in many cases can be explained on the basis of the epidemiological behaviour of each virus [63].

The reconstruction of the phylogeny of the influenza subtype H3N2 shows that the evolution of the virus follows a single track in the long term [37]. The tree has a main trunk with branching points representing drifting strains that become extinct in an approximately one-year period. This type of topology likely results from the continuous emergence of new virus variants that can evade recognition by the immune system through a change in the properties of their antigenic sites. In the case of influenza the renew of susceptibles comes from births of new hosts and from regained susceptibility of previously infected individuals to new influenza variants. The short infectious periods of influenza, together with its high genetic variation, leads to a rapid strain turnover, probably one of the reasons producing the strong seasonality of influenza epidemics. An unresolved question concerns the transmission mode of influenza through respiratory droplets containing just one or a few viral particles. This strong reduction in population size could lead to the fixation of mutations during optimization of the virus inside each particular host. The expected result should be a very high diversity among lineages (infecting different hosts) that is not observed in natural populations of the virus. Mathematical models of influenza epidemics have approached this question by taking into account the balance between the production of new strains and the competition-induced stochastic extinction of existing variants. The results obtained show that, if cross-immunity is the only form of competition between strains, there is an exponential growth of diversity. However, if a short-lived immunity that inhibits reinfection by any new strain during a short time is considered in the model, the probability of explosive diversity growth is strongly reduced [64,126].

A different type of phylogenetic tree is exhibited by viruses that provoke a strong cross-immunity, as is the case of measles. In this case the epidemic cycles arise from repeated exhaustion of susceptible hosts combined with the lifelong immunity elicited by the pathogen. An immune response equally effective against all strains does not promote selection of the best adapted genotypes. Therefore, many strains can coexist with relative frequency, and the topology of the phylogenetic tree is mainly determined by spatial-temporal dynamics [63].

The situation is clearly different in the case of persistent infections, as those established by HIV and hepatitis C virus. The long period of evolution experienced by these viruses inside their hosts means that both dynamics, intra- and inter-host should be reflected in their phylogenetic trees. Due to the long time between transmission events, epidemic cycles are not observable in the case of these viruses. Instead of that, what is observed is a growing epidemic trend. The fact that the infection by a subtype does not protect against new reinfections with other subtypes of HIV or HCV results in a phylogeny that mainly reflects the demographic and spatial history of transmission. The continuous evolution of these viruses inside their hosts permits to perform intra-host phylogenies. In HIV, the highly immunogenic envelope gene presents and intra-host phylogeny similar to that exhibited by influenza A at the population level [63]. The presence of drugs to treat the infection is another factor that strongly conditions the intra-host phylogeny of the virus.

Phylogenies of RNA viruses are very difficult to interpret due to the large number of factors involved. To know the precise interactions between the immune response, the genetic variation of the pathogen and the inter-strain competition (whose intensity depends on the mode of transmission and on the nature of the infection–acute or persistent) does not suffice. Moreover, all these factors must be integrated in the context of population dynamics of the host at the spatial-temporal level. An important advance is being made in this field by contrasting mathematical approaches with

the actual behaviour of viruses in nature. Progress in the amount of field data at hand comes from the large number of virus genome sequences available and from the increased number and precision of epidemiological data.

2.2. *Study of viruses in the laboratory. Experimental protocols*

The laboratory study of viral dynamics in their natural animal hosts is very difficult to carry out. A first obstacle is of practical nature, and refers to the cost of maintaining a pool of susceptible animals. A second difficulty is more fundamental, and concerns the interpretation of the results obtained in a system composed by many different cellular types, not all of them infected by the virus. The interactions with the molecules secreted by the immune system, the ability to spread the infection from cell to cell and other factors that can enormously diverge from an animal to another often mask a clear-cut interpretation of the results. Quite obviously, these studies are strictly necessary when a new vaccine or a new antiviral drug is being investigated, but they are usually preceded by much simpler in vitro studies. In vitro experiments involve "adaptation" to susceptible cells that can be cultivated and that may differ from those infected by the original virus population. Viruses often behave differently on different types of cultured cells, and in addition each of the culture types holds technical advantages and disadvantages. Typically, adaptations to the new conditions in the laboratory involve serial transfers of the virus in liquid cell cultures that provide the appropriate environment for competition among a large number of genomes, eventually leading to the selection of the most fit to the new situation. Usually this process implies acquisition of genetic alterations that confer a reproductive advantage in the new host.

Quantitative virus assays are essential to obtain useful evolutionary data. The plaque assay is the commonest biological assay for viruses. It was developed originally by d'Herelle [22] and adapted to animal viruses by Dulbecco and Vogt [36]. This method enables the cloning of individual genetic variants of a virus and permits to quantify some properties of virus isolated either from natural environments of from liquid cultures. The plaque assay is based on the ability of a single virus to cause cytopathology in a relatively short time interval. Cytopathology is usually identified by an area of lysed cells on an otherwise homogeneous monolayer of cultured cells on a Petri dish. The virus starting the infection produces new viruses that propagate to the surrounding cells. The final result after multiple rounds of infection is what is called a plaque.

Environmental perturbations and changes in the population size stand among the most common alterations with evolutionary consequences. In laboratory conditions, both can be applied in a controlled way to viral evolving populations. Modifications of the population size usually refers to changes affecting the number of viral particles originating a new quasispecies. When this number is strongly reduced, the viral progeny results from a few, even a single, founder particle. In situations where a virus is plated on a cell monolayer using low multiplicity of infection (MOI), only one virion infects a cell, typically. The effect of serial plaque-to-plaque transfers in virus evolution involves isolation of a plaque, extraction of the viral population and, after appropriate dilution, repeated plating on fresh cellular monolayers. Serial plaque-to-plaque transfers enhance the fixation of mutations in the population. In addition, since most mutations in an optimised population are deleterious, fitness loss is the expected result accompanying repeated population bottleneck events. In contrast to plaque-to-plaque transfers, large population passages generally lead to increases in viral fitness [94].

After any evolution experiment, it is common to determine the fitness variations experienced by the virus and the mutations that have been acquired by the viral genomes. These quantifications can be performed after different transfer numbers, allowing in this way to study the degree of divergence from the original population. Fitness values are determined by infecting cells with the virus to be tested together with a genetically or phenotypically marked reference virus.[2] The viral progeny is transferred several times, and the proportion of the two viruses is quantitated at each transfer. In the case of phenotypically distinct viruses, quantitation takes places by plating the virus population in two dishes with and without the selective marker. In the case of genetically different viruses, the genetic material is extracted and the relative amount of each of them is determined by quantitative RT-PCR using specific primers for each virus. A representation of the logarithm of the mutant/ancestral virus as a function of the transfer number typically yields a straight line whose slope is taken as the relative fitness of the evolved population.

To determine the mutations that might be responsible for fitness differences in evolved populations, the consensus sequence is analyzed. As has been explained, only mutations present in most genomes can be detected by conventional

[2] It is common to use as a reference a virus that is resistant to neutralization by a monoclonal antibody.

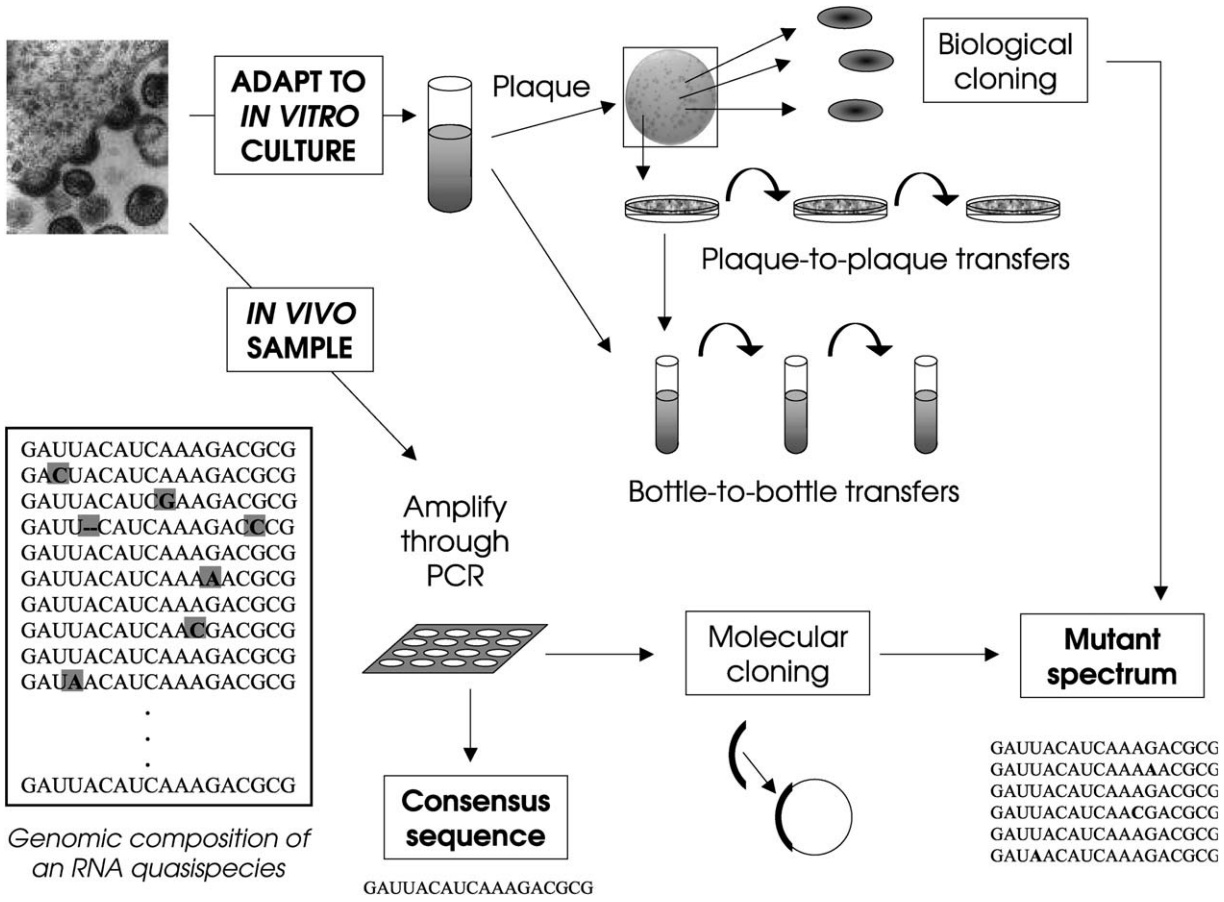

Fig. 2. Scheme of laboratory protocols used to quantify genomic changes in the course of evolution of viral populations. Adaptation from the natural host to the *in vitro* environment is the first necessary step. Isolation of individual virions (for later use as founders of a new population or to characterize the heterogeneity of a quasispecies) is performed through dilution and plating (at low multiplicity of infection) on cellular monolayers deposited on Petri dishes. The virions composing a single lytic plaque constitute a biological clone. Massive passages of a population are usually carried out through bottle-to-bottle transfers. This procedure allows competition between variants and optimization of fitness in the environment considered. At any point, the genomes of the bulk population can be amplified through polymerase chain reaction and sequenced in order to obtain the consensus sequence of the quasispecies (which yields the most common nucleotide at each genomic position). In the case of RNA genomes, translation to DNA through the action of reverse transcriptase, and insertion in the circular genome of a plasmid permits molecular cloning inside a bacteria (*Escherichia coli*, often) and the characterization of population heterogeneity. Since molecular clones are not necessarily viable, while biological clones result from infecting variants, the composition of a quasispecies measured in one or another way might differ. Detailed explanations of these protocols are to be found in Section 2.2.

sequencing methods. The consensus sequence contains part of the information on the evolution followed by the whole quasispecies, and can be of relevance to know the changes experienced by a majority of the genomes. Knowing the diversity in the population necessarily requires sequencing a representative fraction of the ensemble of genomes both in the ancestral and in the evolved viral population. A simple way to do this consists in plating the population on a Petri dish, isolating a number of lytic plaques (called in this case biological clones), and sequencing the genome of the virus population contained in each plaque independently. Since the generation of a plaque involves only a few replication rounds, it is assumed that this sequence is similar enough to that of the virus that initiated the plaque, and thus represents it.

Other genomes, not easily detectable by this method, might exist. They are difficult to detect, either because they are present in a low amount or because they are defective, such that infection of a cell is only effective when they coinfect the same cell with other virion(s) able to provide the proteins that they cannot encode. A technique to detect these defective genomes involves the extraction of the bulk genetic material of the population and cloning in a bacteria such as *Escherichia coli*. In this way one can obtain many recombinant plasmids or molecular clones, including genomes

that are not necessarily infective. Fig. 2 schematically represents the most common protocols used in the laboratory to quantify evolutionary changes in viral quasispecies.

2.3. Viral evolution through population bottlenecks

The experiments to be described in this section were motivated by a theoretical expectation, namely, that population bottlenecks might lead to eventual extinction of a quasispecies due to the steady accumulation of mutations that they promote.

It seems reasonable to assume that external perturbations applied to any optimised population will cause decreases in the average viability of the ensemble. An extreme form of perturbation consists in regularly subjecting the quasispecies to population bottlenecks where a single individual acts as a founder of a new population. On the one hand, the genomes in the new population will, with rare exceptions, carry all of the mutations present in the founder individual. On the other hand, if the development time of the population after the bottleneck is not long enough, advantageous variants will not appear or will be rarely selected. As a consequence, the population will experience a progressive accumulation of mutations that might result in steady degradation and in eventual extinction. This idea derives from theoretical studies carried out by H.J. Muller in the 1960s (see Section 3.2), but has not been experimentally tested until recently.

In the last several years, Escarmís and co-workers have been performing the longest experiment up to date where a virus is subjected to systematic population bottlenecks [28,47–49,74,75]. The RNA virus used as a model system was foot-and-mouth disease virus (FMDV), a pathogen infecting cattle, pigs, sheeps, goats and wild ruminants, among other animals. The disease occurs in most parts of the world, is endemic in many countries, and often causes extensive epidemics. Thus, its study and the ways in which it can be controlled are the subject of current interest. The original quasispecies was well adapted to infect baby hamster kidney cells [117], the host used in the *in vitro* experiments. Four different clones were randomly selected and successively plated on cellular monolayers, following the protocol described in Section 2.2. In those experiments, the number of plaque forming units per ml in the selected lytic plaque was used as a measure of the fitness of the virus. Initially, all the four clones experienced strong decreases in their ability to infect cells and generate abundant and viable progeny. For about 30 to 40 passages, the quasispecies was tolerant to deleterious mutations, such that the number N_{pfu} of infectious units per plaque per ml decreased exponentially with the transfer number t,

$$\log N_{pfu} \propto -ct. \tag{1}$$

Numerical fit to experimental data yield coefficients c between 0.08 and 0.12 [75]. The trend to lose fitness stopped however at about $t \simeq 40$. From then on, the number of viable infective units attained a fixed value on the average, though large fluctuations were observed as population bottlenecks were further applied. The biphasic behaviour observed in those experiments is represented in Fig. 3(a). An analysis of the statistical properties of the fluctuations revealed that their distribution follows a stretched exponential function (Fig. 3(b)). If the values of N_{pfu} are ordered according to their size (such that the largest one has rank $r = 1$, the second largest has $r = 2$ and so on), then the number $N_{pfu}(r)$ follows

$$\left[N_{pfu}(r)\right]^{\alpha} = -a \log r + b, \tag{2}$$

where α is numerically obtained through least squares fit to the experimental data [74,75].

The statistically stationary state is the result of two competing processes. On the one hand, clones of fitness lower than average can be selected at each transfer, such that the average fitness of the population decreases. On the other hand, during the development of a lytic plaque between successive bottlenecks, a certain degree of heterogeneity is generated, both towards lower fitness variants and towards higher fitness variants. As the initial population is placed at a high fitness state, the appearance of less fit mutants is more frequent than the appearance of advantageous ones, as expected. Eventually, however, the generation (and positive selection) of variants of higher fitness balances the appearance (and negative selection) of those with lower fitness. At that point, a statistically stationary state ensues. Seen from another viewpoint, the statistical properties of that state result from the competition of two time scales: that of external perturbations (in the form of population bottlenecks) *vs*. the time allowed for the generation of diversity between bottlenecks. The former time scale is mostly responsible for negative selection and fixation of lower fitness variants, while the latter enhances the appearance of advantageous mutants and enables active positive selection.

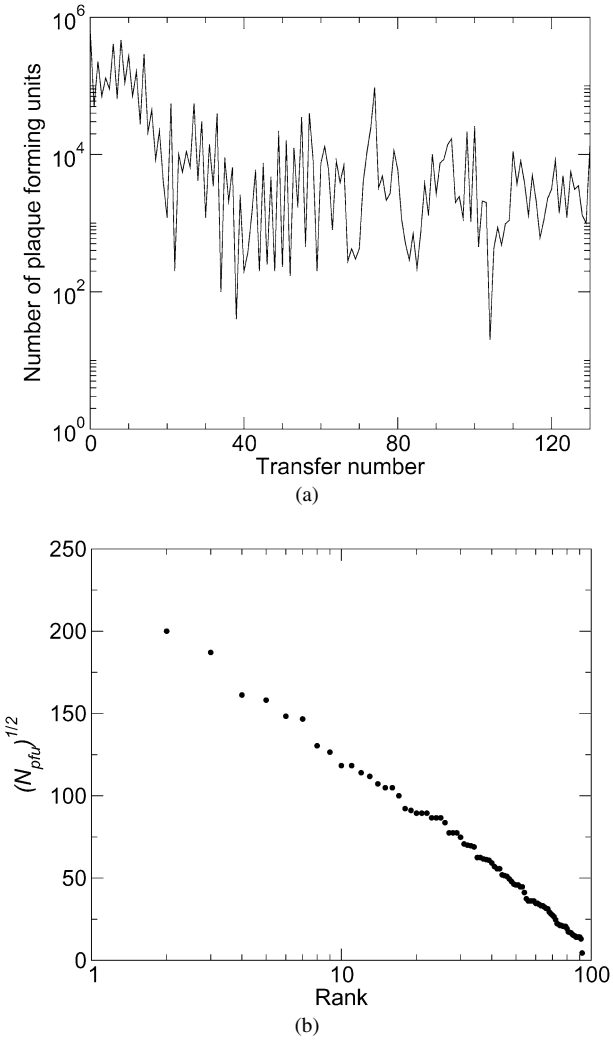

Fig. 3. (a) Number of plaque forming units as a function of the transfer number for one of the FMDV clones analysed by Escarmís and co workers [74]. This experiment applies the most extreme form of population bottleneck to a viral quasispecies, since a single individual acts as a founder of the new population after each transfer. Starting with a well adapted population, fitness decreases as theoretically expected. After about 40 transfers, a statistically stationary state sets in. (b) Rank ordering of the pfu's values from transfer 40 to transfer 130. The fluctuations in this quantity follow a stretched exponential distribution (Eq. (2), with an exponent $\alpha \simeq 0.5$ for all the clones analysed [75].

Indeed, in some cases where lytic plaques could be barely observed or, apparently, infectious particles were absent, an increase in the development time between bottlenecks permitted the recovery of an infective subpopulation [74].

Interestingly, the consensus sequence of the quasispecies accumulated mutations steadily all through the process, at an average pace of one new nucleotide per three transfers [49]. This nonetheless, fitness maintained an average value and occasionally attained high values. Hence, an additional consequence of the empirical observations is that beneficial, or compensatory mutations, cannot be exceedingly rare in this scenario (as opposed to back mutations, which can be discarded on probabilistic grounds). This contrasts with the assumptions of most theoretical models proposed prior to performing those experiments, where the appearance of mutants with fitness higher than that of the fittest type at a given time was discarded (see Section 3.2).

The experiment carried out by the group of Escarmís and Domingo is the longest, but not the first, where RNA viral quasispecies were repeatedly subjected to population bottlenecks. The earliest experiments of this kind go back to 1990, and involved the bacteriophage Φ6 [15]. That phage has a genome divided in three independent segments inside the viral capsid. A significant decrease of infectivity was observed after 40 consecutive transfers which constituted

the whole length of the experiment, but neither a statistically stationary state of fitness values, nor extinction of the population, were observed. Other experiments used vesicular stomatitis virus (VSV) as a model system [35,45]. VSV is a negative-sense, ssRNA virus with a single, non-segmented genome, and does not experience recombination at a detectable level. In those experiments, fitness decreased steadily from the average value of the initial, well-adapted population, in agreement with theoretical expectation. As an additional result [43,95], a dependence of the fitness loss with the initial fitness was observed for the first time: high fitness clones lose fitness faster than clones with lower fitness. Selection pressure is stronger in the latter, such that they are less tolerant to the incorporation of new mutations. Studies with the phage MS2 support the decrease in fitness from an initially well adapted population when bottlenecks are applied [23], with a coefficient $c \simeq 0.16$ (Eq. (1)).

Extinctions of infectivity were not observed, neither with VSV, nor with MS2. This contrasts to the sensitivity of HIV-1 [136], where in a comparable experiment 6 out of 10 clones could not produce viable progeny after only 15 plaque-to-plaque transfers. This observation might reflect a different internal time scale in different virus, in the sense that different viruses employ different time lengths in a replication cycle, or in the killing of a cell. It might be that HIV cannot undergo a sufficiently large number of replication cycles to generate the minimum diversity needed to overcome the negative effects of bottlenecks. Clearly, in the limit where the development time between bottlenecks tends to zero, the probability that an advantageous mutant appears tends to zero as well.

Despite the debilitating effect that the systematic application of bottlenecks exerts upon viral populations, and the accompanying increase in the mutational load, fitness can be recovered in those quasispecies. When the size of the bottleneck is varied, it has been observed that increases or decreases of fitness depend on the mean initial fitness of the population [95]: starting with the same initial fitness, it can increase or decrease depending on the size of the founder population. In order to optimize fitness, large population passages, which permit competition among mutants and positive selection of best adapted variants, are applied. In principle, the ideal mutation-selection equilibrium occurs under unbounded growth of the population. Massive population transfers are thus the experimental procedure that better represents a situation where mutation-selection equilibrium would be reached.

Recovery of fitness in viral populations that had evolved under application of bottlenecks has been experimentally observed [48,96]. Studies of the molecular changes experienced by the recovered populations revealed compensatory mutations in many cases (implying positive epistasis), deletions of fragments of variable size, and occasionally reversions (back mutations) in the consensus sequence. Though the effect on fitness of mutations in the genotype is usually very difficult to assess, the experiments point out to an enormous ability of RNA viruses to find pathways in the genome space able to outweight the negative effects of mutations accumulated through bottlenecks.

2.4. Viral evolution under large population passages

Large population passages are the standard protocol to optimise fitness. Some experiments have studied the process of fitness gain, and have observed that, in general, a first phase of rapid fitness increase is followed by a regime where increase in fitness slows down [94]. A plausible interpretation of this behaviour comes from simple evolutionary models where a first phase dominated by the formation of a quasistationary distribution in fitness space is followed by a slower increase in the average fitness of the population [127]. Interestingly, empirical observations also reveal that the time interval between successive increases in fitness during adaptation follows a stretched exponential distribution [46]. Qualitatively, these studies point out that the population experiences more difficulties in acquiring new advantageous mutations when it is near optimal fitness. We have previously seen that the dynamics of fitness loss through plaque-to-plaque transfers displays an analogous behaviour: a phase of exponential fitness decrease is followed by a statistically stationary state of fitness.

Massive population passages usually imply that many viral particles infect each individual cell. When this happens, selection processes take place at different levels, as well as interference and competition between viral variants, both outside and inside the cell. One form of interference is caused by competition among different viruses for critical replicative pathways, including competition for cell surface receptors (extracellular interference) or competition for biosynthetic machinery (intracellular interference). In other cases, defective (nonmultiplying) viruses can successfully replicate inside the cell if infective forms supplying the proteins that the defective forms lack are concurrently present. Due to the heterogeneity of the quasispecies, a fraction of the population is unavoidably defective, and behaves in practice as a parasite of the infective forms under high multiplicity of infection.

Usually the amount of defective genomes maintains an equilibrium with the amount of fully viable genomes. However, there are situations in which this balance is broken, as for example when there is a large enough increase in the mutation rate [62]. In this case the amount of defective genomes can largely exceed that of the genomes able to encode the correct proteins. This can constitute a form of error catastrophe based on the extinction of the viable population through the lethal action of defectors [62]. Also preextinction RNA obtained from a mutagenized population of FMDV interfere and delay viral production when it is cotransfected with standard RNA [59].

Competition and interference processes represent complex forms of interaction among the genomes constituting a quasispecies. Under low multiplicity of infection (population bottlenecks as described in the previous section) only fully viable genomes can survive. That is, if an isolated virion cannot successfully infect, replicate and encapsidate inside a cell, it will be necessarily eliminated in a bottleneck event. However, under high MOI, defective genomes, unviable in isolation, can thrive due to the simultaneous presence of complete genomes. In fact, defective viral genomes are frequently generated upon serial high multiplicity of infection passage of RNA virus, and have been detected experimentally [111].

2.4.1. Viral sex, coinfection, and competition inside the cell

At present, possible advantages of symbiotic, defective variants can be directly compared to the corresponding standard type. One interesting open question is the nature of subsequent evolutionary steps in the system of defectives, which might help understand the role of viral sex in evolution. If sex is understood as the exchange of genetic material between organisms, then recombination could be interpreted as a form of sex, because fragments of different parental molecules are joined to produce a new genome. However, sex in virus usually refers to forms with segmented genomes that can experience genetic reassortment. This type or genomic reorganization is clearly beneficial for viruses, such as influenza, that can infect different hosts and thus have genes that, in spite of being quite similar, maintain host-specific differences. This kind of sex has clear benefits in the diversification of influenza virus.

The advantages conferred by sex are less clear in other viruses, such as the bacteriophage Φ6, adapted in the laboratory to infect a single host. In this system, the advantage of sex could rely in the fixation in the same genome of mutations occurring in separate lineages. However, when replicate populations of Φ6 were allowed to evolve in either the presence or the absence of sex for a large number of generations, there were no observable differences in the rate of adaptation. Moreover, results indicated that, in some cases, sex was not advantageous because the cost of intrahost competition was too high [124].

Coinfection [87] may be beneficial in large populations of viruses because it permits sexual exchange between viruses. Nonetheless, it may be detrimental because it allows virus complementation, thus allowing inferior genomes to benefit from superior viral products within the cell. The mutational load in the case of Φ6 was purged faster in the absence of coinfection, which suggests that, in this case, the disadvantages of complementation can outweigh the benefit of sex [54].

In experiments in which a wild type virus population of VSV was coinfected with a monoclonal antibody-resistant mutant at different multiplicities of infection, it was observed that the fitness of the mutant showed a great dependence of changes in the multiplicity of infection, which is an indicator of the degree of complementation. When complementation occurs, the fitness of all viruses that coinfect a cell is the average fitness in the absence of coinfection of that group of viruses [97]. There are interesting effects in this same system when situations with low MOI and high MOI alternate. At low MOI only complete genomes can infect and survive, while defectives will survive at high MOI thanks to complementation. The alternation of the two situations promotes coexistence of different viral strains, provided they are alternatively superior in each of the two environments considered [135].

A related phenomenon holds with viral infections in nature. In many cases, the only possibility is infection by a single virus variant. However, if the density of hosts is much lower than the density of viruses, the same host can be coinfected by two or more different variants. Another possibility is that of superinfection [98], where after infection by a virus variant, a second infection by a different variant comes about. In this case the first virus has been evolving inside the host before infection by the second virus occurs. The observations to date show that coexistence by competition during viral evolution is not feasible in the long run. When both coinfecting virus variants possess equal fitness and abundance, they coexist until the fixation of beneficial mutations in a variant provides a large enough advantage to exclude the other [18].

2.4.2. An evolutionary transition towards a multipartite genome

In nature, the wild type of most viruses is represented by a complete genome (in a single piece or in a number of fragments) encapsidated in each virion (see Fig. 2). However, there are some viruses, mostly infecting plants, where two or more defective (incomplete) forms need to enter a cell concurrently in order to infect it. Probably, either form of the genome represents a different adaptive strategy to the natural environment of the virus, and complete or defective forms (the latter requiring coinfection) confer different advantages in different situations [81].

The natural form of foot-and-mouth disease virus (FMDV) consists of a single molecule genome. However, there are other examples of picornavirus-like plant RNA viruses with multipartite genomes. A transition from a complete to a multipartite genome had never been observed until recently, in a long experiment where FMDV was subjected to hundreds of passages at high multiplicity of infection [57]. Indeed, only in a scenario where coinfection with different viral types is not limiting is it possible to observe the appearance and maintenance of several defective forms in the quasispecies. There are, in addition, two theoretical hypothesis proposing that defective, complementary viral forms, might be even able to outcompete the standard form of a virus. One hypothesis is that genome segmentation and encapsidation in different particles evolved in response to high mutation rates, in what would constitute viral sex [16,103]. The second hypothesis states that segmentation could arise if shorter molecules replicate faster enough than full length partners, such that the transition would be favored at high MOI [93,121].

In a recent experiment, an initially fit clone of FMDV [26] was serially passaged at high MOI. After 260 passages, viral RNA was extracted and analysed. Three defective mutants (labeled $\Delta 417$, $\Delta 999$, and $\Delta 1017$), with large deletions in their genomes (of 417, 999, and 1017 nucleotides, respectively), were identified. In isolation, they were unable to infect, but the simultaneous presence of complete and defective forms in the same population maintained high fractions of genomes with deletions. In fact, at passage 260, the standard, complete form of the virus was 10^4 times less abundant than the defective genomes. Two defective, complementary forms were isolated: $\Delta 417$ lacked the L-protease coding region and $\Delta 999$ lacked a part of the capsid-coding region, their deletions occurring on non-overlapping parts of the genome. An interesting observation regarding the virulence of the system of defectives as compared to the standard type was that the former produced lytic plaques remarkably smaller than those of the latter.

A common assay to quantify virulence consists in evaluating the time T required by a virus to completely kill a cellular monolayer, as a function of the initial number n_0 of viral particles infecting cells. This measure was used to compare the two types obtained in the experiments. Interestingly, the functional form of the function $T(n_0)$ varied qualitatively with the viral type. While a dependence of the form $\ln T^d(n_0) \propto \ln(n_0)$ was obtained for the defective type, the killing time follows $T^s(n_0) \propto \ln(n_0)$ for the standard form of the virus [57]. The difference can be explained on the basis of the efficiency of long-range transport in the two-dimensional system where infection propagates. While the standard form can cause infection with a single particle, the requirement of coinfection for the defective forms translates in the inefficiency of long-range transport [85]. In other words, infection of neighboring cells occurs at high MOI, while the dilution caused by long-range transport implies that infection at large distances from the infectious focus occur at low MOI, in which case defection is a losing strategy.

2.5. Effect of mutations on the phenotype

It has been mentioned that one of the consequences of the high error rate of RNA virus replication is that they have a very compact genome. One of the possible side effects of this fact is that, once the quasispecies is highly optimised, almost all new mutations might have an effect on fitness, however small.

Most studies concerning adaptation to environmental changes are focused in the identification of mutations fixed in the consensus sequence with a putative role in fitness changes observed in the virus. Sometimes, these studies have also dealt with the identification of mutations in individual genomes of the quasispecies. This last situation represents a more precise approximation to the real situation holding in a heterogeneous population. Among the environmental perturbations whose effects have been studied one can mention treatments with antiviral drugs, presence of mutagens (which increase the mutation rate of the virus), changes in the population size, changes of tropism, or presence of monoclonal antibodies. All of these perturbations usually translate into a strong decrease in the fitness of the virus, but in most cases the previous presence, or the rapid generation of adapted genomes, favours the acquisition of additional mutations that guarantee survival and drastically modify the composition of the quasispecies. Unfortunately, due to the many different processes involved in successful completion of the viral cycle, it is extremely difficult to assess the effect that a single mutation may have on the viral phenotype.

Occasionally, some studies have determined precisely the effect of a point mutation, a task which might be simplified if the mutation causes a change of aminoacid and eventually induces a measurable change in the codified protein. An example of such a situation was observed in the Vaccinia Virus, a DNA virus where two forms, the standard type and a mutant with a single base change from C to A were compared. The mutation induces a change of aminoacid, from asparagine to alanine-25, and this modification causes a destabilization in the α-helix domain of an envelope protein. The ultimate consequence is the production of smaller lytic plaques in culture, and a reduction in the virulence of the virus [58].

A general observation is that the fraction of deleterious mutants maintained in the population is higher under the systematic application of bottlenecks. In unperturbed populations, negative selection would likely eliminate through competition less fit mutants, while they have a chance to be selected for if bottlenecks are blindly applied (as is the case). In a study with VSV, mutations were found both in coding and non-coding regions of the genome, and non-synonymous mutations were more frequent than synonymous ones [96]. When massive passages were applied in order to recover high fitness states, an interesting observation arose, namely, the quasispecies increased its average fitness without showing obvious changes in the consensus sequence.

In studies with FMDV [47], non-synonymous mutations represented about 45% of the total, and about 50% of the clones acquired a poly-A track with different lengths that turned out to be deleterious for the virus [49]. More detailed studies of the potential effect of mutations on the structure of viral proteins and on fitness of the virus allowed to determine several structural modifications, as changes in nucleotides affecting internal proteins of the capsid, amino acid replacements in non-structural proteins, destabilizations in the secondary structure of the genome, or inactivations in the binding sites of certain proteins, among others. Still, the quantitative determination of the effects of mutations (on FMDV as well as in any other virus) involves cloning of mutants and growth-competition experiments with the standard form. Also of interest are the molecular solutions found by debilitated forms subjected to bottlenecks when massive passages are applied as a means to recover fitness. True reversions are rarely found, while different forms of compensation are common, including shortening of the poly-A internal track and a large deletion of 69 nucleotides including the poly-A section [48].

The effect of mutations on fitness is not independent of the (genetic and environmental) context where the mutations occur [13,131]. A first step towards quantifying the role of the genetic context consists in comparing the effect of a single mutation with the effect of pairs of mutations. This involves the systematic investigation of many mutants, and competition experiments in order to determine their relative advantage. In one such study [104], 62 genotypes of VSV carrying pairs of nucleotide substitutions with respect to a standard type were generated and analysed. In many cases, a significant interaction between mutations was found, since their combined effect was not a simple sum of individual effects. That is to say, epistasis (either antagonistic or synergistic) was present. Other studies have also identified epistasis between mutations. In the case of the phage $\Phi 6$ [14], it was observed that antagonistic epistasis was more frequent the lower the fitness of the viral genome, while major epistatic effects have not been observed with FMDV [44].

As the previous investigations point out, the precise effect of mutations on fitness depends on many different and interwoven factors. There is still a long way until phenotypic variation can be related to molecular changes. Further, the effort to understand the roots of phenotypic change needs to combine the molecular level with intermediate levels including the molecular structure, gene expression, or gene interaction, to name a few. This adds on the role played by the external environment, be it the metabolic state of the cell or the different selective pressures applied on the virus. Altogether, this is what eventually establishes the (relative, context-dependent) evolutionary value of mutations.

2.6. Extinction of viruses through the use of mutagens

The extreme ability of RNA viruses to adapt to changing environments difficults the control of viral diseases. Population bottlenecks were considered as a possible mechanism to increase mutational load and eventually cause extinction. However, the experiments carried out to date support the robustness of viral quasispecies, since systematic extinction cannot be provoked through that procedure. Indeed, in nature it is common that transmission of pathogens involves strong reductions in the population size, this being thus a common perturbation with which they have to cope, and to which they are consequently adapted.

The high plasticity of viruses implies that conventional treatments with drugs become eventually inefficient, due to the appearance of resistance mutants. Further, once they have appeared, those mutants can be present in the

quasispecies even in very small amounts. One consequence of this effect is the mentioned memory of viral quasispecies [109]. Hence, one of the most active fields of research on virology involves seeking mechanisms able to cause the extinction of viral populations (usually, this refers to their infectivity) before they become adapted to the new environment.

The high mutation rate of RNA viruses is partly puzzling. The adaptive advantages of generating a large number of variants, most of them unavoidably unviable, are not completely clear [55]. On the one hand, the repertoire of variants might indeed enhance the appearance of new adaptations [66]. On the other hand, a high mutation rate is risky in that it has to be kept below an upper limit in order to maintain the information coded in the genome [11,39]. Some analyses seem to indicate that the mutation rate of RNA viruses is close to that upper threshold [67,113]. As a result, an interesting antiviral strategy consists in increased mutagenesis [27,30,40], through which it should possible, in principle, to push the quasispecies beyond that threshold. In that case, information would be lost and the population would become unviable.

Though attemps to quantify the adverse effects of chemical mutagenesis are relatively recent, a number of viruses have been already explored, among which VSV and poliovirus [20,67,76], HIV [80], FMDV [99,116], and LCMV [61,62]. In all cases, viral infectivity was strongly reduced and extinction often followed. Another consequence of increased mutagenesis is that the spectrum of mutants broadens: Analyses of the variants present show that the diversity of genomes largely increases, implying that the quasispecies becomes more complex and that the consensus sequence reveals less about the structure of the population. In addition, the fraction of defective genomes increases, and the interaction between viable and parasitic classes acquires higher relevance. In a recent experiment with FMDV, it seems that it is precisely the action of the parasites what drives the population to extinction [62].

Lymphocytic choriomeningitis virus causes persistent infections in baby hamster kidney and Vero cells in the laboratory [61]. Such infections were monitored under the addition of different amounts of the mutagenic base analog 5-fluorouracil. A number of cellular monolayers were infected and, at different times post infection, the total amount of RNA and the total number of infective particles were quantified [62]. A representative plot of the dynamics of the system is shown in Fig. 4. Interestingly, as time elapsed, if any change was to be observed in the intracellular amount of progeny RNA, that was a slight increase. However, the number of plaque forming units decreased, until infectivity disappeared at about 90 hours post infection. The results indicate that the replicative ability of the genome does not disappear concomitantly with infectivity, and suggest the following interpretation of this behaviour. Inside the cell, the products codified by all the infecting viral genomes are available to all of them simultaneously. Since the infection is persistent, cells are not lysed and the main competition among types in the quasispecies is for resources to replicate. The fastest sequences will proportionally increase in number inside the cell. This egoistic behaviour turns out to be lethal for LCMV, since the ability to correctly codify the necessary products for survival is not maintained in the parasitic subpopulation.

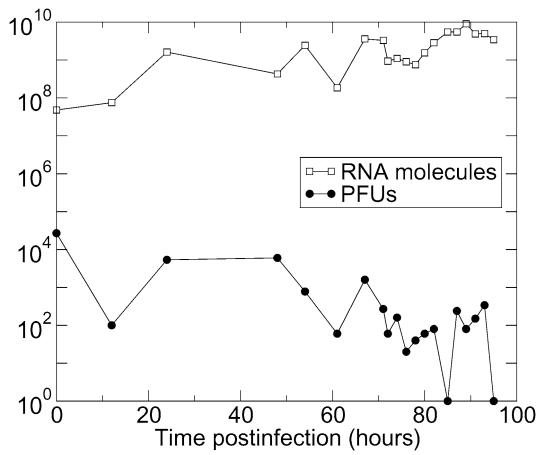

Fig. 4. Changes in intracellular RNA content and number of plaque forming units as infection under the action of a mutagen progresses. Replication is a phenotypic trait that experiences positive selection inside the cell (there is sustained competition between replicating types during a persistent infection). Infectivity is not selected for, since the infection of fresh cells is not necessary for survival. Eventually, the virus loses its infective capacity and becomes unviable, such that extinction follows. In this case, 100 μg/ml of the mutagen 5-fluorouracil were added to the infected cells.

Lethal mutagenesis constitutes a very promising strategy for RNA virus extinction. However, the main obstacle for its effectivity is the appearance of drug resistant mutants [100]. Combined therapies consisting of mutagens and other antivirals could be a more realistic strategy for the treatment of viral infections [1].

3. Theoretical approaches to viral evolution

Many of the concepts currently used to describe and characterize the evolutionary behaviour of quasispecies stem from theoretical analysis of simpler, toy model systems, occasionally able to capture the essential mechanisms behind the observed dynamics. Some of the models that we are going to discuss aimed at describing in a general setting the fate of mutants in populations where mutations appear rarely, more in the spirit of classical population genetics. This is the case of models related to Muller's ratchet, a theoretical concept which eventually inspired the experiments where viral populations are subjected to population bottlenecks. It has been only with the advent of modern techniques, and thanks to the use of viruses as model systems, that some of the frameworks assumed in those models could be recreated in a laboratory.

In its turn, the first theory on self-replication was partly a consequence of in vitro experiments of molecular evolution, as described in the next section. Later, that theory inspired experiments on evolution under increased mutagenesis as a way to induce extinction of viral populations, and in what represents a direct application of the concept of error threshold. The real feedback between theory and experiment has just started, and will for sure pose many challenging questions and yield wonderful and unexpected answers in forthcoming years.

3.1. Molecular quasispecies and the Qβ phage

Molecular replication is acknowledged as a crucial step in the origin and evolution of life. Understanding of the basic mechanisms involved in such process would be possible if extracellular experiments of molecular replication could be performed in a simple and robust setting. At the beginning of 1960's, and after more than half a century of study of bacteriophages, the replicase of phage Qβ turned out to provide a suitable system. It is stable in isolation [118], is a constant environmental factor (not subjected to evolution when replicating RNA in vitro), and accomplishes imperfect replication of RNA strands [10].

The initial experiments carried out by the group of Spiegelman and co-workers [90,119] focused on the discriminative ability of the replicase to accept RNA templates. The initial molecular mixture contained a small amount of RNA from phage Qβ, the replicase, and nucleoside triphosphates at concentration similar to the cellular interior. In those conditions, competition between RNA strains was just for resources and replication: the fastest variants increased exponentially in number and were preferably selected for the next transfer. After 75 transfers, the replication had increased almost three fold, and the length of the replicating RNA had decreased to only 15% the length of the standard Qβ genome. Varying the selective pressures in subsequent experiments revealed a wealth of phenotypic differences in the evolved RNA populations [78].

Interestingly, in another series of experiments with the same system, it was possible to obtain short replicating RNA molecules (up to 50 nucleotides in length) generated in the absence of a template [8,9], provided the concentration of nucleotides and Qβ replicase was high enough. The sequence of the replicating molecules was different in different experiments, except for the terminal nucleotides. However, their secondary structures seemed to be partly analogous. This is also an interesting result and points to the many solutions at the level of sequence yielding similar structure and probably similar molecular function.

In most experiments where replicative ability is the only criterion for selection, replicating molecules evolve shorter sequences. Hence, one might be tempted to conclude that genome length is what determines replication time. However, there are some situations where, contrary to this expectation, replicons evolve longer genomes [86]. One possible interpretation is that the optimal size of replicators reflects a trade-off between the information they encode and the information encoded in the replicating machinery. In any case, this result is a word of caution to infer the behaviour of evolving systems from expectations based on theory or even on indirect experience. Biological systems, and viruses as evolving systems, are a continuous source of unexpected results that require an incessant revision of mechanistic models.

3.1.1. The theory of molecular quasispecies

The results of Spiegelman and co workers stay at the basis of Manfred Eigen's theory of molecular evolution. That theory was not only aimed at explaining the in vitro observations with Qβ replicase, but was also a general theory of molecular evolution at prebiotic stages, when the accuracy in the replication processes was not high enough, and coexistence among a number of different, interrelated types was probably the rule [39]. The theory considers a number $i = 1, \ldots, N$ of different molecular types such that, upon replication, type i can originate type j with a certain probability W_{ij}. Error free replication is given by the diagonal terms W_{ii}. Each molecular type is represented by a fraction x_i of elements in the population, and the evolution takes place inside a flow reactor, such that

$$\sum_{i}^{N} x_i = 1. \tag{3}$$

All those elements yield the basic evolution equation of molecular quasispecies,

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \big(W_{ii} - \overline{E}(t)x_i(t)\big) + \sum_{k\neq i} W_{ik}x_k(t), \tag{4}$$

where $\overline{E}(t)$ is the average excess growth rate of the total population and can be explicitly calculated by imposing the normalization condition (3). At the steady state of evolution, that is, setting $\dot{x}_i = 0$, $\forall i$, the fractions x_i attain constant values. Those values are the elements of the right eigenvector corresponding to the largest eigenvalue of the matrix W_{ij}. The steady state densities define the mutation-selection equilibrium of the quasispecies.

The parameters W_{ii} can be decomposed as the product of two terms: the fidelity parameter Q_{ii} and the growth rate A_i of type i. In turn, the Q_{ii} depend on the length ν_i of the sequence that is being replicated: If each nucleotide of that sequence is faithfully replicated with probability $q_i < 1$, then $Q_{ii} = q_i^{\nu_i}$. In its original formulation, Eigen's theory predicted the existence of an error threshold (a minimum value of q_i) below which a sequence with a selective superiority σ_m would disappear from the quasispecies. Suppose that sequence m, the master sequence, is such that it feeds other classes of mutants when incorrectly replicating, but mutants do not revert to the master class. At the steady state,

$$\bar{x}_m = \frac{W_{mm} - \overline{E}_{k\neq m}}{E_m - \overline{E}_{k\neq m}} \tag{5}$$

in a first approximation. This equation can be written in the form

$$\bar{x}_m = \frac{\sigma_m Q_{mm} - 1}{\sigma_m - 1}, \tag{6}$$

and, in order to maintain $\bar{x}_m \geqslant 0$, one requires

$$1 - q_m \leqslant \frac{\ln \sigma_m}{\nu_m}, \tag{7}$$

for a per nucleotide fidelity $q_m \simeq 1$. Eq. (7) sets a relationship between sequence length, copying fidelity, and selective superiority, and is the core of Eigen's paradox on the limits of information maintainability through selection. If the inequality is violated (the sequence becomes too long, the fidelity is too low, or the superiority is not higher enough than that of the following class) the master sequence disappears from the population.

The initial theory has been enlarged, complemented and refined by a large number of subsequent works. Evolution of finite populations, and thus the effect of population fluctuations on survival of the master sequence, led to a stronger condition on the replication fidelity, though the qualitative results remained unchanged [24]. This additional restriction on the error threshold is however counterbalanced when selection acting on the phenotype is taken into account. Indeed, the many different genotypes that result in the same phenotype (mutational neutrality) [68] translate into an increase in the error threshold when the degeneracy of the map genotype-phenotype is considered. Though this increase is limited, as shown by theoretical studies [122], analysis of real ribozymes suggest that the classical threshold (7) could be increased up to eight fold [73], thus remarkably alleviating Eigen's paradox. This also implies that two different error thresholds could be defined: a first one where the master sequence is lost, and a second one where the phenotypic information is lost. A review of the notion of error threshold can be found in [11].

The structure of the fitness landscape, and its effects on the evolution of two quasispecies of similar selective value was addressed by Schuster and Swetina [112]. They reached the conclusion that increases in the mutation rate could cause a shift in the dominant quasispecies due to differences in the distribution of neutral nearest neighbours in sequence space. This result has been reformulated [133] and applied to computational models, where it has been observed that, under increased mutagenesis, the winner of a competition between two quasispecies is not necessarily the one occupying the highest fitness peak, but may be the population with the largest neutral neighbourhood (thus the most robust one) [132]. Quasispecies do not only maximize fitness through evolution: in order to minimize the effects of mutations in constant environments, they also evolve towards maximal robustness [128]. Some aspects of this phenomenon were already observed in the experiments with Qβ [10], and in subsequent experiments with RNA viruses [13]. Much theoretical interest has been and is currently devoted to the topological properties of the fitness [25] and structure space [52,53], which determine quasispecies robustness and the evolutionary fate of RNA molecular populations [134].

Eigen's theory of molecular quasispecies has been used in different forms to study virus evolution from a theoretical viewpoint. Indeed, this theory is a suitable framework to describe structured, heterogeneous populations, and to address qualitative questions involving competition between quasispecies, interactions with the environment, response to varying parameters, or the effects of increased mutagenesis, among others.

3.2. Muller's ratchet

The appearance of change is an unavoidable effect of reproduction. As a consequence, and if positive selection is limited, deleterious mutations necessarily accumulate unless they are eliminated by sex. The increase in the mutational load was studied, among others, by Muller [92], who compared the unavoidable accumulation of mutations in asexual populations to the clicks of a ratchet mechanism. The experiments described in Section 2.3 intend to decrease fitness through the frequent application of population bottlenecks, thus equating the accumulation of mutations to degradation of the population. This inference, however, was not made by Muller in his original work [50].

In asexual populations, deleterious mutations can be fixed, particularly if the population size is small [21]. Eventually, an asexual population of size N subjected to a mutation rate μ and a selection coefficient s against deleterious mutations attains a deterministic mutation-selection equilibrium. At that point, there is a fixed amount of genomes n_m with m mutations. The equilibrium distribution of the population was calculated by Kimura and Maruyama [70] and reads

$$n_m = N\frac{\exp(-\mu/s)}{m!}\left(\frac{\mu}{s}\right)^m. \tag{8}$$

As long as N is large enough, there will be a significant amount of wild-type (mutation free) genomes,

$$n_0 = N\exp(-\mu/s). \tag{9}$$

However, if N is small or the coefficient μ/s is sufficiently large, the number of genomes in the mutation-free class can be small enough to be affected by population fluctuations that might eventually cause the disappearance of all of the individuals in that class. If this happens, the ratchet has clicked once and the least loaded class corresponds now to individuals carrying at least one mutation. The one-mutation class can disappear in the same manner, and the ratchet clicks again. This model implicitly assumes that reversions are the only mechanism able to produce an increase in fitness. If the highest fitness class is identified with the mutation-free class and is assigned value 1, the fitness $w(m)$ of individuals with m mutations will be

$$w(m) = (1-s)^m. \tag{10}$$

This defines a multiplicative fitness landscape for the mutated genomes and implicitly assumes that each mutation affects fitness independently (epistatic effects are not included).

Nowadays, most interpretations of Muller's ratchet equate accumulation of mutations with fitness loss. The rationale behind this identification comes from presupposing that the mean genotypic value of the population is close to the optimum. The first formal study of Muller's ratchet under the previous assumptions, and using Eq. (8) was carried out by Haigh [65] on a model proposed by Felsenstein [51]. He concluded that the speed of the ratchet (the time between two successive clicks or losses of the least-loaded class) depended chiefly on n_0. However, more recent

analyses [60,120] reveal that the speed of the ratchet depends not only on the number of individuals in the least mutated class, but also on the parameters μ and s, meaning, as could have been expected, that it is the whole structure of the population, the quasispecies, that determines the fate of each and every class of mutants present.

Other authors analysed and discussed different properties of Muller's ratchet under similar assumptions and proposed recombination as the most effective mechanism in order to arrest the ratchet and thus recover high-fitness (or mutation free) genomes [71,89]. Epistatic effects among mutations where also considered as a possible mechanism able of accelerating or slowing down the speed of the ratchet. Under weak antagonistic epistasis [17] the time between successive clicks of the ratchet grows, though the rate of fitness decline is barely affected (as in previous models, it is assumed that all mutations have a negative effect on fitness). However, under stronger epistatic selection the ratchet can be effectively halted, such that, in practice, a finite population could survive almost indefinitely [72].

3.2.1. Population bottlenecks

Few models have explicitly considered the introduction of population bottlenecks, since this process introduces a stochastic component more difficult to deal with formally. In [19], the development of the population between bottlenecks was assumed to be long enough such that mutation-selection equilibrium distribution could be attained. Further, a fitness function of the form

$$w(m) = (1 - s)^{m^{\alpha}} \tag{11}$$

was considered. The parameter α is the epistasis parameter: synergistic and antagonistic epistasis are described by $\alpha > 1$ and $\alpha < 1$, respectively, and Eq. (10) (absence of epistasis) is recovered for $\alpha = 1$. While, in all cases studied, antagonistic epistasis slows down the ratchet and synergistic epistasis accelerates it, other effects, such as a decrease of the population size, facilitated by the steady accumulation of mutations, can notably accelerate the extinction of the population due to mutational meltdown [82].

Earlier attempts to include population bottlenecks in mathematical models were motivated by the form of transmission of mitochondrial genomes in mammals, which have no recombination and mostly undergo monoparental inheritance. With a simple model devised to represent such a situation, Bergstrom and Pritchard [6] demonstrated that bottlenecks may be essential in maintaining mitochondrial genetic quality, in contrast with previous theoretical expectations, which predicted an easier accumulation of mutations in small populations.

However, the large amount of asexual, non-recombining species subjected to strong population bottlenecks in their natural environments which, this nonetheless, seemed to keep a high average fitness, turned the attention to other forms in which the continued degeneration predicted by most models could be compensated. The occurrence of compensatory mutations (not reversions) which could have phenotypic effects and even outweigh the effect of the accumulated deleterious mutations has been considered [2,7,74,108,129]. Indeed, compensatory mutations were proposed as a mechanism as effective as recombination in halting the deleterious effect of the ratchet [129], in the framework of a model that includes as a novelty the separation of effects in the phenotype from changes in the genotype. This distinction is especially relevant in the context of quasispecies, since the mutation-selection equilibrium is defined as the stationary distribution of fitness values selected in the given environment. This equilibrium does not fix the consensus sequence, which would surely be affected by (quasi-)neutral drift. Starting with a high-fitness variant, decreases in fitness due to the action of the ratchet were observed. However, at a certain point compensatory mutations stop the effect of deleterious mutations and the average fitness of the population keeps an average value with small fluctuations around it [129]. Hence, the dynamics display a biphasic behaviour qualitatively analogous to the empirical observations [74,75]. Actually, all models where compensatory mutations are explicitly included permit to recover higher fitness states in the population, and to escape extinction. As a side effect, biphasic evolution is observed whenever the initial fitness of the population is far from the average stationary fitness at the corresponding statistically stationary state. This is so even when bottlenecks, and thus strong fluctuations in the population size, are systematically applied. As we know now empirically, compensatory mutations are indeed one of the main mechanisms to maintain high fitness.

Aiming at understanding the relationship between the mode of transmission of a pathogen and its virulence, a simple model where the population was structured in several fitness classes was proposed [7]. A down-mutation (occurring with probability p) moved the offspring of a genome from fitness class w to class $w - 1$, while an up-mutation increased fitness from w to $w + 1$ with probability q. By means of numerical simulations, it was shown that vertically transmitted pathogens suffer decreases in fitness much stronger than when transmission is horizontal. A similar model

was numerically [74] and analytically [83] studied in order to understand the appearance of biphasic behaviour and, in particular, of the large fluctuations in the viral yield described in Section 2.3. Additional parameters of this model are the size of bottlenecks and the development time between their application [84]. Increases in the number of individuals going through the bottleneck implies that the average fitness at the statistically stationary state increases, as experimentally observed [95]. The development time determines the distance to the mutation-selection equilibrium and constrains the degree of heterogeneity of the population before the bottleneck.

In fact, the equilibrium distribution in Eq. (8) is attained after a long development time, for large enough populations, and only if all mutations have the same effect on fitness. Most populations, even in carefully controlled experiments, are expected to be out of the mutation-selection equilibrium. In particular, bottlenecks condition the structure of the population and its phenotypic characteristics in non-trivial ways. Following a bottleneck, both the frequency and variance of the population's genetic composition change [137]. If bottlenecks are regularly applied such that the development time of the population between bottlenecks is much shorter than the time required to achieve mutation-selection equilibrium, the (out-of-equilibrium) composition of the population $P(w)$ in the presence of deleterious and beneficial mutations can be analytically calculated [83],

$$P(w) \propto \left(\frac{q}{p}\right)^{w-1} \left(\frac{w+1}{2}\right)^{g} \tag{12}$$

where g is the (integer) number of replication cycles that take place between two consecutive bottlenecks.[3] Eventually, this function permits the explicit calculation of the distribution of the viral yield, which in this model follows a stretched exponential function, in qualitative agreement with empirical observations.

4. Final remarks

Most of the experiments carried out to date consider adaptation of viral populations, well adapted to a certain environment of origin, to a new one. Commonly, the variables that define each of the environments are kept constant, so no environmental fluctuations are introduced through evolution. In such conditions, once the population is optimised, what would be the advantages conferred by the continuous generation of new variants? This addresses a new question of evolutionary relevance in the long run, namely, the selection of optimal mutation rates as a function of the environmental changes that the population is exposed to [110].

In nature, there are always environmental novelties, situations not encountered before, that require a certain degree of plasticity in the population to escape extinction. This issue has been tackled recently from a theoretical point of view [38], with the conclusion that rapid or dramatic environmental change indeed promotes the selection of higher mutation rates, while static environments have as a consequence a decrease in the mutation rate. Though quantitative changes in the mutation rate might be difficult to detect empirically due to the large time scales involved, certain observations hint at the fact that an increase in the mutation rate can be promoted by large enough, random environmental changes [101,106,115].

Viruses are hierarchically structured and complex dynamical systems where evolution and selection act at different levels and with broadly different timescales. They are the result of a long evolutionary process along which many different strategies for survival have been explored. Empirical results bring about surprises very often, and force to reconsider not only the simplifying assumptions of some theoretical models, but the precise meaning of concepts related to quasispecies evolution.

There is a strong need of continuous dialog between theory and experiment. In the near future, the interplay between the two approaches looks promising. On the one hand, computational abilities already permit to perform detailed simulations of evolving populations where phenotype and genotype are separated properties. RNA folding algorithms, for instance, are accurate enough to address evolutionary questions where the feature under selection is the molecular structure, not the sequence, as assumed in most computational models up to now. On the other hand, the amount of empirical results steadily increases, and in vitro techniques enable the analysis of single molecules in

[3] Since both deleterious and compensatory mutations are considered in this model, the fitness class w includes individuals with a varying number of mutations, depending on their history. Thus, the relationship between $P(w)$ and n_m, as in Eq. (8) is not immediate. The introduction of compensatory mutations breaks the equality between less mutated class and fittest class.

evolving populations. Analytical approaches become more difficult as more details are implemented in the description of the system, though simple models that can be analytically studied may result in the introduction of new concepts and can be of aid in the isolation of the relevant (universal) mechanisms at play. Though, often, they cannot be directly applied to predict the outcome of real experiments, they might determine what is possible under a set of simple rules. Their failure has to translate into their improvement. Computational approaches may act as a bridge between analytical and experimental studies, since they can embrace more detailed descriptions of evolving systems and are extremely relevant to educate our guesses on real virus evolution.

The conjunction of analytical, computational, and experimental in vitro studies will hopefully result in unprecedented advances in our understanding of the evolutionary mechanisms acting at the molecular level.

Acknowledgements

The authors acknowledge fruitful and continued discussions with friends and colleagues: U. Bastolla, C. Biebricher, C. Briones, E. Domingo, C. Escarmís, J. García-Arriaza, A. Grande-Pérez, J. Pérez-Mercader, and the support of INTA and Ministerio de Educación y Ciencia (FIS2004-06414). SCM benefits from a Ramón y Cajal contract.

References

[1] Arnold JJ, Vignuzzi M, Stone JK, Andino R, Cameron CE. Remote-site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. J Biol Chem 2005;280:25706–16.
[2] Bachtrog D, Gordo I. Adaptive evolution of asexual populations under Muller's ratchet. Evolution 2004;58:1403–13.
[3] Ball LA. Replication strategies of RNA viruses. In: Fields virology. 4th ed. New York: Raven Press; 2001.
[4] Baranowski E, Ruiz-Jarabo CM, Domingo E. Evolution of cell recognition by viruses. Science 2001;292:1102–5.
[5] Batschelet E, Domingo E, Weissmann C. The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. Gene 1976;1:27–32.
[6] Bergstrom CT, Pritchard JK. Germline bottlenecks and the evolutionary maintenance of mitochondrial genomes. Genetics 1998;149:2135–46.
[7] Bergstrom CT, McElhany P, Real LA. Transmission bottlenecks as determinants of virulence in rapidly evolving pathogens. Proc Nat Acad Sci 1999;96:5095–100.
[8] Biebricher CK, Eigen M, Luce F. Template-free RNA synthesis by Qβ replicase. Nature 1986;321:89–91.
[9] Biebricher CK, Luce R. Sequence-analysis of RNA species synthesized by Qβ replicase without template. Biochem 1993;32:4848–54.
[10] Biebricher CK, Gardiner WC. Molecular evolution of RNA in vitro. Biophys Chem 1997;66:179–92.
[11] Biebricher CK, Eigen M. The error threshold. Vir Res 2005;107:117–27.
[12] Boerlijst MC, Bonhoeffer S, Nowak MA. Viral quasispecies and recombination. Proc R Soc Lond B Biol Sci 1996;263:1577–84.
[13] Burch CL, Chao L. Evolvability of an RNA virus is determined by its mutational neighbourhood. Nature 2000;406:625–8.
[14] Burch CL, Chao L. Epistasis and its relationship to canalization in the RNA virus Φ6. Genetics 2004;167:559–67.
[15] Chao L. Fitness of RNA virus decreased by Muller's ratchet. Nature 1990;348:454–5.
[16] Chao L. Levels of selection, evolution of sex in RNA viruses, and the origin of life. J Theor Biol 1991;153:229–46.
[17] Charlesworth D, Morgan MT, Charlesworth B. Mutation accumulation in finite outbreeding and inbreeding populations. Genet Res 1993;61:39–56.
[18] Clarke DK, Duarte EA, Elena SF, Moya A, Domingo E, Holland J. The Red Queen reigns in the kingdom of RNA viruses. Proc Nat Acad Sci 1994;91:4821–4.
[19] Colato A, Fontanari JF. Soluble model for the accumulation of mutations in asexual populations. Phys Rev Lett 2001;87:238102.
[20] Crotty S, Cameron CE, Andino R. RNA virus error catastrophe: Direct molecular test by using ribavirin. Proc Nat Acad Sci 2001;98:6895–900.
[21] Crow JF, Kimura M. An introduction to population genetics theory. New York: Harper and Row; 1970.
[22] d'Herelle F. The bacteriophage and its behavior. Baltimore, MD: Williams and Wilkins; 1926.
[23] de la Peña M, Elena SF, Moya A. Effect of deleterious mutation-accumulation of the fitness of RNA bacteriophage MS2. Evolution 2000;54:686–91.
[24] Demetrius L, Schuster P, Sigmund K. Polynucleotide evolution and branching processes. Bull Math Biol 1985;47:239–62.
[25] Derrida B, Peliti L. Evolution in a flat fitness landscape. Bull Math Biol 1991;53:355–82.
[26] Domingo E, Dávila M, Ortín J. Nucleotide sequence heterogeneity of the RNA from a natural population of foot-and-mouth disease virus. Gene 1980;11:333–46.
[27] Domingo E. Antiviral strategy on the horizon. Vir Res 2005;107:115–6.
[28] Domingo E, Escarmís C, Lázaro E, Manrubia SC. Quasispecies dynamics and RNA virus extinction. Vir Res 2005;107:129–39.
[29] Domingo E, Holland JJ. RNA virus mutations and fitness for survival. Annu Rev Microbiol 1997;51:151–78.
[30] Domingo E, Pariente N, Airaksinen A, González-López C, Sierra S, Herrera M, Grande-Pérez A, Lowenstein PR, Manrubia SC, Lázaro E, Escarmís C. Foot-and-mouth disease virus evolution: Exploring pathways towards virus extinction. Curr Topics Microbiol Immunol 2005;288:149–73.

[31] Domingo E, Ruiz-Jarabo CM, Sierra S, Arias A, Pariente N, Baranowski E, Escarmís C. Emergence and selection of RNA virus variants: memory and extinction. Virus Res 2002;82:39–44.
[32] Domingo E, Sabo DL, Taniguchi T, Weissmann C. Nucleotide sequence heterogeneity of an RNA phage population. Cell 1978;13:735–44.
[33] Drake JW, Charlesworth B, Charlesworth D, Crow JF. Rates of spontaneous mutation. Genetics 1998;148:1667–86.
[34] Drake JW, Holland JJ. Mutation rates among RNA viruses. Proc Nat Acad Sci USA 1999;96:13910–3.
[35] Duarte E, Clarke D, Moya A, Domingo E, Holland J. Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc Nat Acad Sci 1992;89:6015–9.
[36] Dulbecco R, Vogt M. Some problems of animal virology as studied by the plaque technique. Cold Spring Harbor Symp Quant Biol 1953;18:273–9.
[37] Earn DJD, Dushoff J, Levin SA. Ecology and evolution of the flu. TRENDS in Ecol and Evol 2002;17:334–40.
[38] Earl DJ, Deem MW. Evolvability is a selectable trait. Proc Nat Acad Sci 2004;101:11531–6.
[39] Eigen M. Self-organisation of matter and the evolution of biological macromolecules. Naturwissenschaften 1971;58:465–523.
[40] Eigen M. Error catastrophe and antiviral strategy. Proc Nat Acad Sci 2002;99:13374–6.
[41] Eigen M, Biebricher CK. Sequence space and quasispecies distribution. In: Domingo E, Holland JJ, Ahlquist P, editors. RNA genetics: Variability of RNA genomes, vol III. 1st ed. Boca Raton, FL: CRC Press; 1988. p. 211–45.
[42] Eigen M, Schuster P. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. Naturwissenschaften 1977;64:541–65.
[43] Elena SF, González-Candelas F, Novella IS, Duarte EA, Clarke DK, Domingo E, Holland JJ, Moya A. Evolution of fitness in experimental populations of vesicular stomatitis virus. Genetics 1996;142:673–9.
[44] Elena SF. Little evidence for synergism among deleterious mutations in a nonsegmented RNA virus. J Mol Evol 1999;49:703–7.
[45] Elena SF, Moya A. Rate of deleterious mutations and the distribution of its effects on fitness in vesicular stomatitis virus. J Evol Biol 1999;12:1078–88.
[46] Elena SF, Sanjuán R. RNA viruses as complex adaptive systems. BioSyst 2005;81:31–41.
[47] Escarmís C, Dávila M, Charpentier N, Bracho A, Moya A, Domingo E. Genetic lesions associated with Muller's ratchet in an RNA virus. J Mol Biol 1996;264:255–67.
[48] Escarmís C, Dávila M, Domingo E. Multiple molecular pathways for fitness recovery of an RNA virus debilitated by operation of Muller's ratchet. J Mol Biol 1999;28:495–505.
[49] Escarmís C, Gómez-Mariano G, Dávila M, Lázaro E, Domingo E. Resistance to extinction of low fitness virus subjected to plaque-to-plaque transfers: Diversification by mutation clustering. J Mol Biol 2002;315:647–61.
[50] Escarmís C, Lázaro E, Manrubia SC. Population bottlenecks in quasispecies dynamics. Curr Topics Microbiol Immunol 2005;299:141–70.
[51] Felsenstein J. Evolutionary advantage of recombination. Genetics 1974;78:737–56.
[52] Fontana W, Schuster P. Continuity in evolution: On the nature of transitions. Science 1998;280:1451–5.
[53] Fontana W, Schuster P. Shaping space: the possible and the attainable in RNA genotype-phenotype mapping. J Theor Biol 1998;194:491–515.
[54] Froissart R, Wilke CO, Montville R, Remold SK, Chao L, Turner PE. Co-infection weakens selection against epistatic mutations in RNA viruses. Genetics 2004;168:9–19.
[55] Furió V, Moya A, Sanjuán R. The cost of replication fidelity in an RNA virus. Proc Nat Acad Sci 2005;192:10233–7.
[56] Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, Cummins LB, Arthur LO, Peeters M, Shaw GM, Sharp PM, Hahn BH. Origin of HIV-1 in the chimpanzee Pan troglodytes. Nature 1999;397:436–41.
[57] García-Arriaza J, Manrubia SC, Toja M, Domingo E, Escarmís C. Evolutionary transition towards defective RNAs that are infectious by complementation. J Virol 2004;78:11678–85.
[58] Gong S, Lai C, Dallo S, Esteban M. A single point mutation of Ala-25 to Asp in the 14,000-M_r envelope protein of Vaccinia Virus induces a size change that leads to the small plaque size phenotype of the virus. J Virol 1989;63:4507–14.
[59] González-López C, Arias A, Pariente N, Gómez-Mariano G, Domingo E. Preextinction viral RNA can interfere with infectivity. J Virol 2004;78:3319–24.
[60] Gordo I, Charlesworth B. The degeneration of asexual haploid populations and the speed of Muller's ratchet. Genetics 2000;154:1379–87.
[61] Grande-Pérez A, Sierra S, Castro MG, Domingo E, Lowenstein PR. 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 Nat Acad Sci 2002;99:12938–43.
[62] Grande-Pérez A, Lázaro E, Lowenstein P, Domingo E, Manrubia SC. Suppression of viral infectivity through lethal defection. Proc Nat Acad Sci 2005;102:4448–52.
[63] Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, Holmes EC. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 2004;303:327–32.
[64] Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of influenza evolution. Nature 2003;422:428–33.
[65] Haigh J. Accumulation of deleterious genes in a population—Muller's ratchet. Theor Pop Biol 1978;14:251–67.
[66] Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, Vandepol S. Rapid evolution of RNA genomes. Science 1982;215:1577–85.
[67] Holland JJ, Domingo E, de la Torre JC, Steinhauer DA. Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. J Virol 1990;64:3960–2.
[68] Huynen MA, Stadler PF, Fontana W. Smoothness within ruggedness: the role of neutrality in adaptation. Proc Nat Acad Sci 1996;93:397–401.
[69] Jensen JH. Isolation of yellow-mosaic virus from plants infected with tobacco mosaic. Phytopathology 1933;23:964–74.
[70] Kimura M, Maruyama T. The mutational load with epistatic gene interactions in fitness. Genetics 1966;54:1337–51.

[71] Kondrashov AS. Selection against harmful mutations in large sexual and asexual populations. Gen Res 1982;40:325–32.
[72] Kondrashov AS. Muller's ratchet under epistatic selection. Genetics 1994;136:1469–73.
[73] Kun Á, Santos M, Szathmáry E. Real ribozymes suggest a relaxed error threshold. Nat Gen 2005;37:1008–11.
[74] Lázaro E, Escarmís C, Domingo E, Manrubia SC. Modeling viral genome fitness evolution associated with serial bottleneck events: Evidence of stationary states of fitness. J Virol 2002;76:8675–81.
[75] Lázaro E, Escarmís C, Pérez-Mercader J, Manrubia SC, Domingo E. Resistance of virus to extinction on bottleneck passages: Study of a decaying and fluctuating pattern of fitness loss. Proc Nat Acad Sci 2003;100:10830–5.
[76] Lee CH, Gilbertson DL, Novella IS, Huerta R, Domingo E, Holland JJ. Negative effects of chemical mutagenesis on the adaptive behavior of vesicular stomatitis virus. J Virol 1997;71:3636–40.
[77] Levine AJ. The origins of virology. In: Fields virology. 4th ed. New York: Raven Press; 2001.
[78] Levisohn R, Spiegelman S. Further extracellular Darwinian experiments with replicating RNA molecules–Diverse variants isolated under different selective conditions. Proc Nat Acad Sci 1696;63:805–7.
[79] Loeffler F, Frosch P. Berichte der Kommission zur Erforschung der Maul- und Klauenseuche bei dem Institut für Infectionskrankheiten. Part I, 23:371–391. In: Milestones in microbiology: 1556 to 1940. ASM Press; 1998. p. 149–53. Translated and edited by Thomas D. Brock.
[80] Loeb LA, Essigmann JM, Kazazi F, Zhang J, Rose KD, Mullins JI. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. Proc Nat Acad Sci 1999;96:1492–7.
[81] López-Ferber M, Simón O, Williams T, Caballero P. Defective or effective? Mutualistic interactions between virus genotypes. Proc R Soc Lond B 2003;270:2249–55.
[82] Lynch M, Gabriel W. Mutation load and the survival of small populations. Evolution 1990;44:1725–37.
[83] Manrubia SC, Lázaro E, Pérez-Mercader J, Escarmís C, Domingo E. Fitness distribution in exponentially growing asexual populations. Phys Rev Lett 2003;90:188102.
[84] Manrubia SC, Escarmís C, Domingo E, Lázaro E. High mutation rates, bottlenecks, and robustness of RNA viral quasispecies. Gene 2005;347:273–82.
[85] Manrubia SC, García-Arriaza J, Domingo E, Escarmís C. Long-range transport and universality classes in *in vitro* viral infection spread. Preprint 2005.
[86] Marshall KA, Ellington AD. Molecular parasites that evolve longer genomes. J Mol Evol 1999;49:656–63.
[87] May RM, Nowak MA. Coinf and evolution of parasite virulence. Proc R Soc London B 1995;261:209–15.
[88] Mc Kinney HH. Factors affecting the properties of a virus. Phytopathology 1926;16:753–8.
[89] Maynard SJ. The evolution of sex. Cambridge: Cambridge University Press; 1978.
[90] Mills DR, Peterson RL, Spiegelman S. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. Proc Nat Acad Sci 1967;58:217–24.
[91] Moya A, Holmes EC, González-Candelas F. The population genetics and evolutionary epidemiology of RNA viruses. Nat Rev Microbiol 2004;2:279–88.
[92] Muller HJ. The relation of recombination to mutational advance. Mut Res 1964;1:2–9.
[93] Nee S. The evolution of multicompartmental genomes in viruses. J Mol Evol 1987;25:277–81.
[94] Novella IS, Duarte EA, Elena SF, Moya A, Domingo E, Holland JJ. Exponential increases of RNA virus fitness during large population transmissions. Proc Nat Acad Sci 1995;92:5841–4.
[95] Novella IS, Elena SF, Moya A, Domingo E, Holland JJ. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. J Virol 1995;69:2869–72.
[96] Novella IS, Ebendick-Corpus BE. Molecular basis of fitness loss and fitness recovery in vesicular stomatitis virus. J Mol Biol 2004;342:1423–30.
[97] Novella IS, Reissig DD, Wilke CO. Density-dependent selection in vesicular stomatitis virus. J Virol 2004;78:5799–804.
[98] Nowak MA, May RM. Superinfection and the evolution of parasite virulence. Proc R Soc London B 1994;255:81–9.
[99] Pariente N, Sierra S, Lowenstein PR, Domingo E. Efficient virus extinction by combinations of a mutagen and antiviral inhibitors. J Virol 2001;75:9723–30.
[100] Pfeiffer JK, Kirkegaard K. Ribavirin resistance in hepatitis C virus replicon-containing cell lines conferred by changes in the cell line or mutations in the replicon RNA. J Virol 2005;79:2346–55.
[101] Plotkin JB, Dushoff J. Codon bias and frequency-dependent selection on the hemagglutinin epitopes of influenza A virus. Proc Nat Acad Sci 2003;100:7152–7.
[102] Posada D, Crandall KA, Holmes EC. Recombination in evolutionary genomics. Annu Rev Genet 2002;36:75–97.
[103] Pressing J, Reanney DC. Divided genomes and intrinsic noise. J Mol Evol 1984;20:135–46.
[104] Sanjuan R, Moya A, Elena SF. The contribution of epistasis to the architecture of fitness in an RNA virus. Proc Nat Acad Sci 2004;101:15376–9.
[105] Rhodes T, Wargo H, Hu WS. High rates of human immunodeficiency virus type 1 recombination: near-random segregation of markers one kilobase apart in one round of viral replication. J Virol 2003;77:11193–200.
[106] Rosenberg SM. Evolving responsively: adaptive mutation. Nature Rev Genetics 2001;2:504–15.
[107] Rota PA, Oberste MS, Monroe SS, Nix W, Campagnoli R, Icenogle JP, Penaranda S, Bankamp B, Maher K, Chen MH, Tong S, Tamin A, Lowe L, Frace M, DeRisi JL, Chen Q, Wang D, Erdman DD, Peret TC, Burns C, Ksiazek TG, Rollin PE, Sanchez A, Liffick S, Holloway B, Limor J, McCaustland K, Olsen-Rasmussen M, Fouchier R, Gunther S, Osterhaus AD, Drosten C, Pallansch MA, Anderson LJ, Bellini WJ. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science 2003;300:394–9.
[108] Rouzine IM, Wakeley J, Coffin JM. The solitary wave of asexual evolution. Proc Nat Acad Sci 2003;100:587–92.
[109] Ruiz-Jarabo CM, Arias A, Molina-Paris C, Briones C, Baranowski E, Escarmís C, Domingo E. Duration and fitness dependence of quasispecies memory. J Mol Biol 2002;315:285–96.

[110] Sasaki A, Nowak MA. Mutation landscapes. J Theor Biol 2003;224:241–7.
[111] Schlesinger S. The generation and amplification of defective interfering RNAs. In: Domingo E, Holland JJ, Ahlquist P, editors. RNA genetics: Variability of RNA genomes, vol. III. 1st ed. Boca Raton, FL: CRC Press; 1988. p. 167–85.
[112] Schuster P, Swetina J. Stationary mutant distributions and evolutionary optimisation. Bull Math Biol 1988;50:635–60.
[113] Schuster P. RNA based evolutionary optimization. Origins Life Evol Biosph 1993;23:373–91.
[114] Shafer RW. Genotypic testing for human immunodeficiency virus type 1 drug resistance. Clin Microbiol Rev 2002;15:247–77.
[115] Shapiro JA. Genome organization, natural genetic engineering and adaptive mutation. Trends in Genetics 1997;13:98–104.
[116] Sierra S, Dávila M, Lowenstein PR, Domingo E. Response of foot-and-mouth disease virus to increased mutagenesis. Influence of viral load and fitness in loss of infectivity. J Virol 2000;74:8316–23.
[117] Sobrino F, Dávila M, Ortín J, Domingo E. Multiple genetic variants arise in the course of replication of foot-and-mouth disease virus in cell culture. Virol 1983;128:310–8.
[118] Spiegelman S, Haruna I, Holland IB, Beaudreau G, Mills DR. Synthesis of a self-propagating and infectious nucleic acid with a purified enzyme. Proc Nat Acad Sci 1965;54:919–27.
[119] Spiegelman S, Mills DR, Peterson RL. Extracellular evolution of a self-duplicating nucleic acid molecule. Science 1967;156:542.
[120] Stephan W, Chao L, Smale J. The advance of Muller's ratchet in a haploid asexual population: approximate solution based on diffusion theory. Genet Res 1993;61:225–32.
[121] Szathmáry E. Natural selection and dynamical coexistence of defective and complementing virus segments. J Theor Biol 1992;157:383–406.
[122] Takeuchi N, Poorthuis PH, Hogeweg P. Phenotypic error threshold; additivity and epistasis in RNA evolution. BMC Evol Biol 2005;5:9.
[123] Tanaka MM, Bergstrom CT, Levin BR. The evolution of mutator genes in bacterial populations: the roles of environmental change and timing. Genetics 2003;164:843–54.
[124] Turner PE, Chao L. Sex and the evolution of intrahost competition in RNA virus Φ6. Genetics 1998;150:523–32.
[125] de la Torre JC, Holland JJ. RNA virus quasispecies populations can suppress vastly superior mutant progeny. J Virol 1990;64:6278–81.
[126] Tria F, Lässig M, Peliti L, Franz S. A minimal stochastic model for influenza evolution. J Stat Mech 2005:07008.
[127] Tsimring LS, Levine H, Kessler DA. RNA virus evolution via a fitness-space model. Phys Rev Lett 1996;76:4440–3.
[128] van Nimwegen E, Crutchfield JP, Huynen M. Neutral evolution of mutational robustness. Proc Nat Acad Sci 1999;96:9716–20.
[129] Wagner GP, Gabriel W. Quantitative variation in finite parthenogenetic populations: what stops Muller's ratchet in the absence of recombination? Evolution 1990;44:715–31.
[130] Webster RG, Hulse DJ. Microbial adaptation and change: avian influenza. Rev Sci Tech 2004;23:453–65.
[131] Weinreich DM, Watson RA, Chao L. Perspective: Sign epistasis and genetic constraint on evolutionary trajectories. Evolution 2005;59:1165–74.
[132] Wilke CO. Selection for fitness vs. selection for robustness in RNA secondary structure folding. Evolution 2001;55:2412–20.
[133] Wilke CO, Wang JL, Ofria C, Lenski RE, Adami C. Evolution of digital organisms at high mutation rates leads to survival of the flattest. Nature 2001;412:331–3.
[134] Wilke CO, Adami C. Evolution of mutational robustness. Mut Res 2003;522:3–11.
[135] Wilke CO, Reissig DD, Novella IS. Replication at periodically changing multiplicity of infection promotes stable coexistence of competing viral populations. Evolution 2004;58:900–5.
[136] Yuste E, Sánchez-Palomino S, Casado C, Domingo E, López-Galíndez C. Drastic fitness loss in human immunodeficiency virus type 1 upon serial bottleneck events. J Virol 1999;73:2745–51.
[137] Zhang X-S, Wang J, Hill WG. Redistribution of gene frequency and changes of genetic variation following a bottleneck in population size. Genetics 2004;167:1475–92.