ANNUAL REVIEWS

Annual Review of Virology

Historical Perspective on the Discovery of the Quasispecies Concept

Esteban Domingo,[1,2] Carlos García-Crespo,[1] and Celia Perales[1,2,3]

[1]Department of Interactions with the Environment, Centro de Biología Molecular Severo Ochoa (CBMSO), Consejo Superior de Investigaciones Científicas (CSIC), 28049 Madrid, Spain; email: edomingo@cbm.csic.es

[2]Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd) del Instituto de Salud Carlos III, 28029 Madrid, Spain

[3]Department of Clinical Microbiology, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), 28040 Madrid, Spain

Annu. Rev. Virol. 2021. 8:51–72

The *Annual Review of Virology* is online at virology.annualreviews.org

https://doi.org/10.1146/annurev-virology-091919-105900

Keywords

mutation, population dynamics, adaptation, mutation rate, sequence space, residue conservation, antiviral strategies, lethal mutagenesis

Abstract

Viral quasispecies are dynamic distributions of nonidentical but closely related mutant and recombinant viral genomes subjected to a continuous process of genetic variation, competition, and selection that may act as a unit of selection. The quasispecies concept owes its theoretical origins to a model for the origin of life as a collection of mutant RNA replicators. Independently, experimental evidence for the quasispecies concept was obtained from sampling of bacteriophage clones, which revealed that the viral populations consisted of many mutant genomes whose frequency varied with time of replication. Similar findings were made in animal and plant RNA viruses. Quasispecies became a theoretical framework to understand viral population dynamics and adaptability. The evidence came at a time when mutations were considered rare events in genetics, a perception that was to change dramatically in subsequent decades. Indeed, viral quasispecies was the conceptual forefront of a remarkable degree of biological diversity, now evident for cell populations and organisms, not only for viruses. Quasispecies dynamics unveiled complexities in the behavior of viral populations,

with consequences for disease mechanisms and control strategies. This review addresses the origin of the quasispecies concept, its major implications on both viral evolution and antiviral strategies, and current and future prospects.

1. INTRODUCTION

In virology, quasispecies is defined as the set of mutant genomes that comprise viral populations. Mutant ensembles are subjected to episodes of competition, selection, and random drift; they can act collectively as a unit of selection. The concept applies to all RNA viruses for which genomic sequences within infected cells and organisms have been compared. It does not refer to differences among consensus sequences of independent viral isolates, which recapitulate influences on later stages of virus evolution.

The quasispecies population structure was identified during the 1970s and early 1980s, prior to the era of routine sequencing, molecular cloning, and PCR, due to a patient screening for mutations by RNA fingerprinting of biological clones of viruses. Years later, viral genetic heterogeneity was observed by use of rapid sequencing of biological and molecular clones. In the past decade, the extreme complexity of RNA viral populations and their change in mutant composition with time have been fully confirmed by deep sequencing methodologies. Although the rate of genetic change is 10^4- to 10^7-fold larger for RNA viruses than for their host organisms (1), host environments also present nonuniform challenges to viruses. The biosphere diversity in which viruses are installed has consequences for the understanding of virus-host interactions because variant virus forms often meet with unique intracellular environments. The present-day scenario is in sharp contrast with concepts of mutational stasis that prevailed in genetics well into the twentieth century. Because the origin of quasispecies lies in replication with frequent production of mutant genomes, it is also valid for DNA viruses whose replication is catalyzed by low-fidelity polymerases or that use RNA as a replicative intermediate.

Following the discovery of quasispecies, some authors attributed viral genome heterogeneity to sequencing artifacts and even questioned the value of the concept for viral genetics. However, the sequencing of genomes extracted from biological clones (not dependent on in vitro amplification of viral RNA) and quantification controls of the mutations identified in molecular clones rendered it unlikely that high mutation rates and frequencies were overestimates. The levels of genetic heterogeneity inferred from deep sequencing data are reasonably close to those calculated from the initial sequence comparisons, although recent sequencing platforms allow an unprecedented penetration into the repertoire of minority mutations present in viral populations.

Mutant distributions offer an interpretation of adaptation based on the dynamic replacement of genome subpopulations within replicating ensembles. This is not a minor facet because responding to harsh selective constraints or to subtle environmental changes is what viruses have to do most of the time, including when they spread within organisms to cause disease. Interestingly, quasispecies dynamics has provided a molecular interpretation of why some vaccine and antiviral therapeutics fail and has suggested new approaches for viral disease prevention and treatment. Among the latter is lethal mutagenesis—virus extinction by an externally induced excess of mutations—which has acquired new impetus with the need to respond to the coronavirus disease 2019 (COVID-19) emergence (Sections 9 and 10).

The quasispecies concept had two independent origins, one from theoretical biophysics and the other from experimental virology; both took place during the 1970s. Here we succinctly describe them from a historical perspective, with references to present-day virology. At the time of their formulation, however, some connections between theory and empiric observations were

not obvious. John Holland and colleagues (2) were the first to bring to light the biological significance of high mutation rates of RNA viruses in the context of a DNA-based biosphere, with medical implications of the collision between the contrasting DNA and RNA worlds.

Following the account of the origin of quasispecies, we recapitulate its major implications for virology as we perceive them half a century later.

2. ORIGIN OF QUASISPECIES THEORY

Manfred Eigen and Peter Schuster in Göttingen, Germany, developed quasispecies theory to explain self-organization and evolution of primitive RNA (or RNA-like) molecules (replicons) that might have populated Earth at the onset of life. They approached self-organization of replicating molecules by linking principles of Darwinian evolution and information theory (3, 4). Darwinian evolution was explicitly investigated in the experiments of Sol Spiegelman and colleagues (5), who showed that replicating RNA molecules evolved in the test tube. Information theory had previously highlighted the requirement for the biological meaning imprinted in a molecule to be faithfully transmitted to the following generations. The major departure from this traditional view was to describe mathematically a system involving multiplication of molecules with a regular production of error copies as inherent to the replicative process (3, 4) (**Figure 1**). With equation 1 in **Figure 1**, replication and mutation became firmly linked. The error copies of a master sequence (defined as the one with the highest replicative fitness in relation to its error copies) was referred to as a comet tail by Eigen and was later known as mutant spectrum, distribution, cloud, or swarm in virology.

There are two impediments to mutant distributions attaining a population equilibrium implied by the initial formulation of quasispecies theory: (*a*) the random nature of mutations and (*b*) the statistical fluctuations in genome frequencies during the replicative process, which continue upon successive infectious cycles. Mutations are random because they arise from quantum mechanical uncertainties regarding the electronic distribution of some atoms in template and nucleotide substrate residues; altered electronic distributions may modify base pairing or base stacking interactions that exert an influence on template copying. When and where a mutation will arise during RNA virus replication is largely unpredictable; the certainty is that mutations will occur at an average rate of about one mutation per 10,000 nucleotides copied (Section 3). Statistical fluctuations may affect the variant genome production, competition, and selection, mainly through viral population migrations entailing reductions in population size, notably bottlenecks of different intensity (**Figure 1**).

The random nature of mutations (even accounting for differences in mutation type preferences among polymerases or biases at some template sites) concurrent with modifications of virus population size are key ingredients of the unpredictable course of virus evolution. Remarkably, however, features reminiscent of determinism—consisting of similar behavior in parallel evolutionary lineages—are occasionally recorded under some experimental conditions, and they beg for an interpretation at the molecular level. In viral dynamics, a population equilibrium (meaning a steady mutant distribution that is constant in time) is beyond reach under usual replication conditions, even after extensive virus multiplication in cell culture in the absence of external influences. Populations can at most be portrayed as consisting of sequential, very short equilibrium steps (Section 7).

A prediction of quasispecies theory is that there is a limit to the information content that can be stably inherited for a given average template copying fidelity of the replication machinery. This concept is mathematically formulated as an error threshold relationship that includes as one of its terms the average polymerase error rate above which the information cannot be propagated

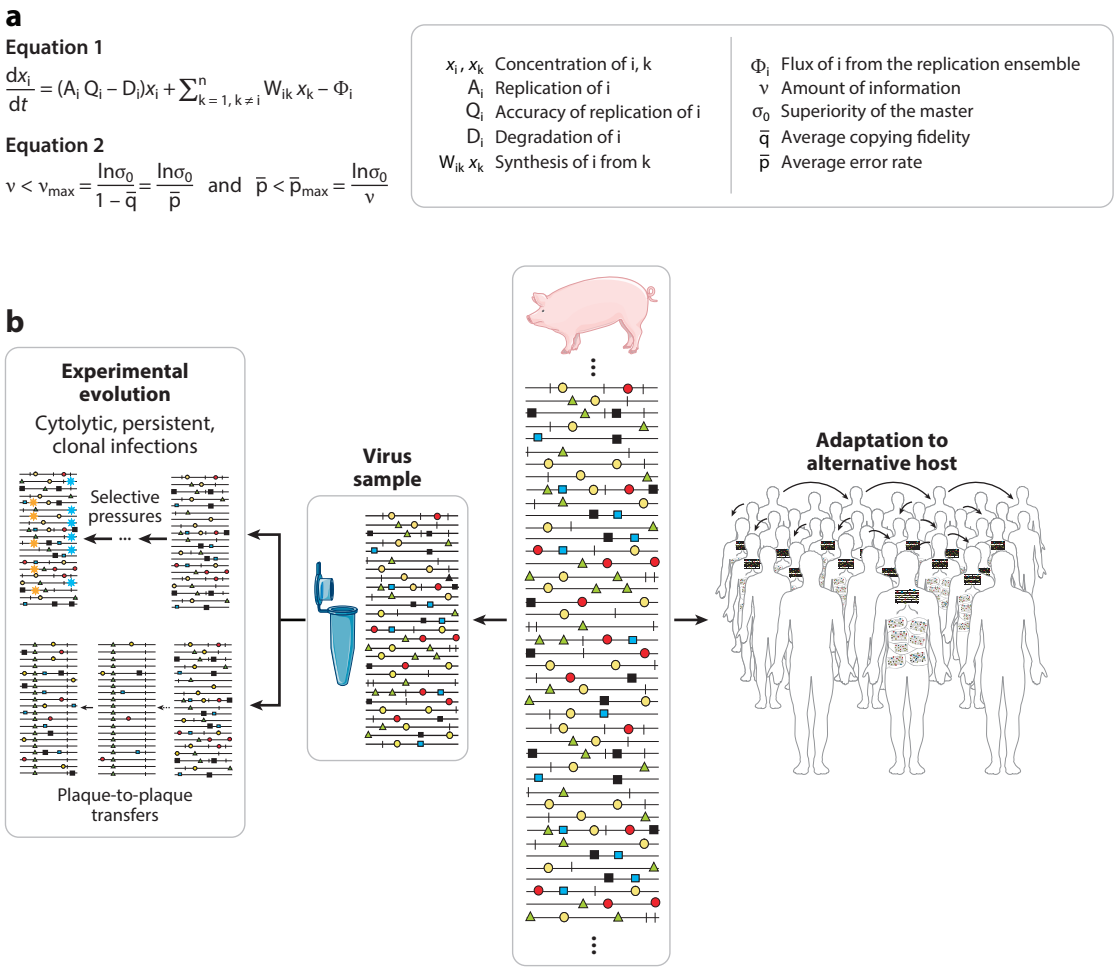

Figure 1

Basic equations of quasispecies theory and meaning of quasispecies for viral populations. (*a*) Equations describe replication (variation of the concentration of genome i as a function of time) with production of error copies (equation 1) and the maximum accepted amount of genetic information that can be stably maintained for a given superiority of the master sequence and average copying fidelity (equation 2). Terms are explained in the box. Derivation of the equations and extensions to nonequilibrium conditions can be found in Reference 8. (*b*) Organisms (in this case a pig) infected by RNA viruses (or some rapidly mutating DNA viruses) include complex and dynamic mixtures of variant genomes (represented by *horizontal lines* with colored symbols to depict mutations). A virus sample is a minimal representation of the variants present in the infected organism, and it may be used for experimental evolution studies to test the effect of selective pressures or bottlenecks, realized by plaque-to-plaque transfers (*left panels*). Quasispecies may facilitate transmission to a new host, in this case humans, where they will form new, diverse, and compartmentalized mutant spectra (*right*). Figure adapted with permission from References 33 and 98.

(equation 2 of **Figure 1**). For viral genomes the information content can be equated with the number of nonredundant protein-coding regions and regulatory elements they harbor. Two transformations can be distinguished: (*a*) the loss of superiority of the master sequence while drift of components of the mutant spectrum is still allowed and (*b*) the complete disappearance of the population due to transgression of an extinction threshold (6, 7). Maintenance of viral identity is compatible with transmission of variant versions of a virus, sometimes exhibiting

different phenotypes. The application of the concept embodied in equation 2 is the basis of lethal mutagenesis as an antiviral strategy (Section 9).

The generally nondeterministic nature of virus evolution does not invalidate quasispecies as a theoretical framework to approach viral dynamics. If the initial theory was not a sufficient stimulus for experimental inquiries, extensions of quasispecies and its associated error threshold relationship have been formulated for entities that exhibit a finite population size and that confront variable fitness landscapes (treated in several chapters of Reference 8).

The quasispecies display of primitive replicons was an important component of the catalytic hypercycle, which was put forward as a mechanism of self-organization in the process of life construction; the hypercycle combines the coding capacity of nucleic acids with the catalytic potential of proteins in sustainment of primitive life forms (4, 9). Before we describe the experimental origin of quasispecies, it is pertinent to mention that Eigen asserted on several occasions that the value of a theoretical model lies in its capacity to evoke experimental testing (3, 10, 11). The experiments that quasispecies theory has inspired in virology have been (and continue to be) highly revealing of the nature of viral populations and of how they adjust to changing environments.

3. EXPERIMENTAL ORIGIN OF QUASISPECIES

The first evidence of RNA virus quasispecies was obtained in the laboratory of Charles Weissmann in Zürich, Switzerland, also in the 1970s, independent of the theoretical development of the concept. The discovery was preceded by many years of studies on the mechanism of *Escherichia coli* bacteriophage Qβ replication, purification of its RNA replicase, and the elaboration of a cell-free system for sustained phage Qβ RNA amplification that even today stands as the most efficient in vitro RNA replication system (12–14).

A stepwise in vitro Qβ RNA synthesis procedure using purified replicase allowed the template-directed introduction of a mutagenic nucleotide analog into preselected positions of the full-size complementary RNA product. In the second round of copying, the analog directed the incorporation of either the correct or the incorrect nucleotide, generating a transition mutation at the preselected site of some progeny genomes. Curiously, the analog N^4-hydroxy-CTP used at the time (15, 16) is the same pyrimidine analog produced intracellularly by prodrug NHC (β-D-N^4-hydroxycytidine-5′-isopropyl ester) used half a century later for lethal mutagenesis of coronaviruses (17, 18) (Section 10). Relevant tools were provided by Walter Fiers and colleagues (19) in Ghent, Belgium, with the initial RNase-based RNA sequencing procedures and a T1-oligonucleotide fingerprinting method that was applied to the identification of mutations in P^{32}-labeled viral RNA (20).

Synthesis of Qβ RNAs carrying a mutation at preselected genomic sites was the birth of site-directed mutagenesis and reverse genetics (21), amply used today in virology and cell biology. A mutant Qβ RNA synthesized by this procedure that included a transition at the 3′ extracistronic region was viable (produced infectious progeny that maintained the engineered mutation). At the time, the result came as a surprise because the high identity of the 3′ extracistronic region among related bacteriophages (that was revealed by the very first results of RNA sequencing) had been taken as evidence that extracistronic mutations would be lethal. The naïve distinction between viability and lethality, without consideration of an intermediate behavior, reflects how our perception of the effect of mutations in viral genomes has changed since then. Upon replication in *E. coli* cells, the extracistronic mutant reverted to the wild-type sequence, indicating that the mutant displayed a selective disadvantage relative to its wild-type counterpart. A rate of 10^{-4} reversion events per nucleotide copied was calculated by combining reversion and wild-type mutant competition experiments (22). The value obtained was several orders of magnitude higher than the

mutation rates calculated by John Drake (23) for some DNA viruses and microorganisms. In the following decades, high mutation rates for RNA viruses and the difference with cells and complex (large genome size) DNA viruses were corroborated (24–26).

Confirmation of the specificity of the site-directed mutagenesis procedure required many infections of *E. coli* spheroplasts with the designed Qβ mutant RNAs and analyses of progeny RNA. In the course of such experiments, it became obvious that phage populations (either the progeny of site-directed mutants or standard, nonmutagenized clonal populations) were composed of pools of mutants present at different frequencies. "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" (27, p. 735). Weissmann presented these results to Eigen and his colleagues at a Max Planck Winter Seminar in Klosters, Switzerland, in January 1978, and the connection between quasispecies theory and experimental observations was made. One of the attendees recalled that Eigen stood up and said, "Quasispecies in reality!" The subject was discussed at several Winter Seminars in following years with participation of Christof Biebricher, whose work with colleagues on in vitro Qβ RNA replication was an essential link between theory and experimentally determined kinetic parameters (28).

The calculation of a mutation rate and the conclusion that phage Qβ populations consisted of mutant spectra provided a quantitative interpretation of many observations with RNA viruses that had been reported in preceding decades, such as frequent reversion of temperature-sensitive mutants and high frequency of plaque morphology bacteriophage variants or lesion-type mutants in plant viruses, among others (additional examples and references are collected in Reference 29). Quasispecies research continued on several theoretical and experimental fronts in addition to virology. These domains include evolutionary optimization and the origin of life, replicator networks, extensions to cellular biology such as the dynamics of bacterial populations and cancer cells, chromosomal instability, conformation heterogeneity of prions, or generalization of the error threshold concept (different areas of quasispecies research have been summarized in Reference 8). Thus, from both theoretical and experimental perspectives, the quasispecies concept has been instrumental for the investigation of dynamic systems, whether they are viruses, cells, or molecules.

4. OVERVIEW OF QUASISPECIES IMPACT FOR VIROLOGY

The introduction of quasispecies in virology led to some controversy because some experts argued that the concept was unnecessary, that it provided a misleading view of virus evolution, and that some virologists referred to the term quasispecies inappropriately (see, for example, the exchange of letters in References 30, 31). Additional concerns were expressed on the theory and the experimental evidence of mutant spectra. It was argued that a deterministic mathematical formulation such as that used for the initial quasispecies theory (implying infinite steady-state mutant distributions in equilibrium) could not represent viral populations. From the experimental side, it was suggested that mutation frequencies might have been overestimated due to errors introduced during viral RNA template amplification procedures (discussed in Reference 1). In our view, four developments in the past two decades have largely resolved the above concerns and have reinforced quasispecies theory as a framework to understand key events in virus-host interaction, viral pathogenesis, and the very early stages of virus diversification and evolution. The first developments are extensions of quasispecies theory to finite populations in variable fitness landscapes—i.e., extensions to mathematically describe mutant spectra under nonequilibrium conditions (1, 8) (see also Section 2). The second developments are observations that suggest that mutant swarms can act as units of selection. They include suppressive effects of mutant spectra on high fitness genomes,

interference, cooperation, or complementation among components of a mutant spectrum, and quasispecies memory. These interactions extend mutant spectrum composition and behavior beyond the standard mutation-selection balance (for reviews of mutant spectra acting as units of selection, see References 1, 8, 32, 33). Third is that confirmation through control experiments that the number of mutations introduced during reverse transcription–PCR amplification procedures prior to molecular cloning and Sanger sequencing and by ultra-deep sequencing did not inflate in a significant manner the number of mutations genuinely present in the RNA template molecules of the virus samples. No significant discrepancies have been observed in quantifications of mutant spectrum complexities using these different methods, or those resulting from sequencing biological viral clones, without cloning of individual molecules amplified in vitro. Fourth is that the large proportion of low-frequency variants and their dynamic change in frequency (e.g., by comparing sequential samples from a virus replicating in cell culture or in a live host organism) currently being revealed by ultra-deep sequencing portray a level of complexity not contemplated in the concept of genetic polymorphism of population genetics, at least as classically formulated (1). Therefore, although alternative procedures are available to understand short-term and long-term viral evolution, quasispecies theory stands as a solid framework to interpret virus-host interactions, as well as virus diversification in individual infected hosts, prior to transmission and subsequent interhost evolutionary events.

Three major domains of quasispecies influence, and some of their ramifications, are schematically represented in **Figure 2**. They are conceptual departures such as a new definition of wild type, the role of mutant spectra in virus adaptability and behavior, and the need to consider viral dynamics when planning antiviral interventions. Central to the ramifications is that

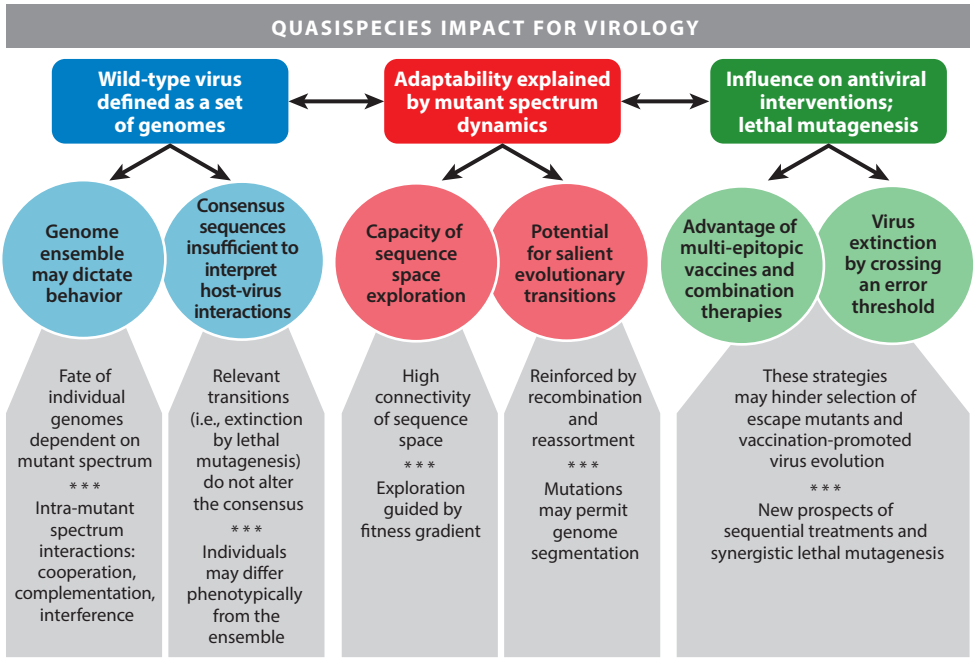

Figure 2

Scheme of quasispecies implications for viruses, and their connections. Major departures are a new meaning of wild type, a more precise understanding of adaptation at the molecular level, and the necessity to open new avenues for the control of viral disease. Each major departure has ramifications written succinctly below the arrows. It puts into focus the main points discussed in this article. See text for evidence and references.

complex mutant spectra—fueled by high mutation rates—are dynamically renewed in response to environmental perturbations and stochastic (chance) effects. The perturbations comprise subtle modifications of environmental parameters (intracellular ionic conditions or metabolite concentrations, temperature, etc.), components of the immune response, and external introductions (antiviral agents, vaccination). Despite our consideration of viral genomes in terms of consensus sequences (a largely unavoidable simplification), a deeper understanding of the nature of viral populations and virus-host interactions is gained from the composition of mutant spectra. The latter cannot be regarded as mere mutant aggregates acting independently of one another; internal interactions of cooperation or interference can be established in what has been called the social behavior of viral populations. The first evidence was provided by the suppressive potential of a vesicular stomatitis virus (VSV) population (32), and additional cases have been reported (reviewed in References 33, 34). The presence in mutant spectra of a molecular memory of those genomes that were dominant in previous phases of the same evolutionary lineage is yet another consequence of the weight of the mutant ensemble, with medical implications (35).

High mutation rates and the quantity of tolerated mutations recorded under various biological circumstances render the presence of mutant spectra the norm; even viruses within a plaque on a cell monolayer are not genetically homogeneous (36, 37). DNA viruses whose replicative machinery displays higher fidelity or includes proofreading-repair activities, and even more those whose progeny DNA can benefit from postreplicative repair pathways, are expected to have a low net mutational input. Their quasispecies features would be displayed under different time frames and population size parameters (38), as also documented for cellular organisms (several chapters of Reference 8).

5. MUTANT SPECTRA AS PHENOTYPIC RESERVOIRS

Mutant clouds equip viral populations with multiple related minority genomes present at low frequency. Those that are better suited to respond to a selective constraint than the dominant ones will increase in frequency once the constraint is in place. Clouds participate in a permanent game of challenges and responses. Basal constraints are those inherent to the cellular environments in which viruses replicate, and they act on both viral protein and RNA (extracistronic but also within open reading frames); RNA structures may be an important part of the viral phenotype (recent example in Reference 39).

In differentiated hosts, clouds vary among organs, tissues, and probably even individual cells (40). The long-term historical adaptation of viruses to the cell types they infect has shaped the core genome organization and encoded proteins to provide replicability and an acceptable long-term survival probability. The central viral genome organization—which determines virus identity and assignment to one of the groups defined by the International Committee on Taxonomy of Viruses—should be flexible enough to face host alterations, including nutritional influences (41). The core genome concept has parallels with that of the pan-genome of bacteria, wherein all members of a bacterial species share a core element while individuals differ in some nonidentical features of their genetic information (42).

Because a large proportion of mutations that occur during viral genome replication results in fitness decrease, mutant spectra tend to be populated mainly by genomes with a limited number of mutations relative to either the nonmutated class (when such a class can be specified) or genomes with the lowest number of mutations in that population (27, 43). Heavily mutated genomes, whose origin has sometimes been traced to apolipoprotein B messenger RNA editing complex, and adenosine deaminase acting on double-stranded RNA activities tend to be rare. They may survive in specific environments—such as the human brain for hypermutated measles virus

RNA derivatives responsible for some neurological syndromes—or due to positive epistatic interactions among mutations or, more generally, tolerance to constellations of accumulated mutations.

Mutations that bring about a phenotypic change have been identified as minority components of mutant spectra of populations that exhibited a different phenotype. Phenotypes embodied in minority genomes cannot be disregarded on the grounds that their becoming prominent is unlikely. A population bottleneck (either intrahost or interhost) entails a certain probability that one of the hitherto hidden traits may be expressed by the chance promotion of the genome encoding that trait to dominance.

The biological effect of a single mutation can be recognized following site-directed mutagenesis of a viral genome (Section 3), but its presence in mutant spectra may not have been directly demonstrated. Its fitness cost may mean it is present at frequencies below detectability. Detectability of a mutation may also depend on the accompanying mutations in the same genome and other genomes. Fortunately, advances in deep sequencing, with increasingly lower cutoff values for mutant detection and new methods for sequencing entire viral genomes (44), should contribute to clarifying if a residue responsible for one phenotype is present at low frequency in the mutant spectrum of the virus exhibiting a different (or the opposite) phenotype.

Still another category of mutations comprises those whose presence in a mutant spectrum is inferred from the experimental design that identifies them. It includes mutants that are resistant to inhibitors, antibodies, or cytotoxic T cells (CTLs) and that are selected from a viral population that collectively displays sensitivity to the constraint. In a study with human immunodeficiency virus type 1 (HIV-1), the magnitude and genetic barrier for CTL escape were comparable to those measured for escape to antiretroviral therapy (45). Escape mutations may incur a fitness cost, but high mutation rates and large population sizes facilitate finding mutational pathways for fitness recovery. The frequency of monoclonal antibody-resistant mutants (MARMs), determined for several RNA viruses, falls in the range of 10^{-3} to 10^{-7}, independent of the number of circulating serotypes of the virus in nature (reviewed in Reference 1). There is no reason to suspect that the mutation rate at the genomic residues that lead to the MARM phenotype differs on average from the mutation rate at other sites for which we lack a quantification marker; the corresponding mutant frequencies may, however, be modified following subsequent replication rounds. These considerations predict the presence in mutant spectra of individual viral genomes with deviant phenotypes, as increasingly reported in the literature.

Selection for an alternative viral phenotype may take place indirectly on infected cells. In parainfluenza virus 5 (PIV5), some amino acid substitutions in protein P—which is part of the virus replication complex—may repress viral transcription and replication, thus allowing subsets of infected cells that express low levels of viral proteins to escape killing by the cellular immune response (particularly by CTLs); such selective survival of P protein mutants mediates PIV5 persistence (46). Both lytic and persistent variants were detected in the mutant spectrum of a dog PIV5 isolate (that had been minimally passaged in cell culture), suggesting the coexistence of lytic- and persistent-prone genomes in PIV5 quasispecies. In general, the presence in mutant spectra of minority genomes exhibiting decreased cytopathology may favor their selection as part of the multiple mechanisms involved in viral persistence (47). Thus, CTL escape can be mediated by amino acid substitutions that impede recognition of the relevant epitopes or indirectly through limited expression of viral proteins at the surface of infected cells.

Quasispecies components may also modulate the immune response. Natural VSV isolates include clones that differ from the majority of the population in capacity to respond to or to induce interferon (IFN), a fact that established IFN induction as a quasispecies marker (48). The IFN response may be affected by amino acid substitutions in viral proteins such as Sendai virus protein C (49) or PIV5 protein V (50). Avian influenza virus subpopulations with enhanced

Table 1 Influence of single mutations on phenotypic traits

Modification(s)	Example(s) and comment(s)	Reference
Lethality	Amino acid substitution at an antigenic determinant of rabies virus glycoprotein that altered virus lethality for mice	112
Establishment of viral persistence	Amino acid change in the lymphocytic choriomeningitis virus glycoprotein that suppresses the antiviral CTL response	113
Lytic and leukemogenic activity	Amino acid substitution in the envelope of Friend murine leukemia virus that attenuated the lytic and leukemogenic properties of this retrovirus	114
Diabetogenicity	Amino acid substitution that renders encephalomyocarditis virus diabetogenic	115
Alteration of neurovirulence	Reduction of neuroinvasiveness associated with neuroinvasiveness	116
Modification or expansion of host range	A substitution in foot-and-mouth disease virus adaptation to guinea pig; substitutions in the IV polymerase	117
Breakage of host barrier	SIV Vif substitution to permit SIV transmission from chimpanzees to gorillas, overcoming the APOBEC barrier	118
Change in receptor and coreceptor recognition and cell tropism	HIV-1 transition from use of CCR5 to CXCR4 as coreceptor; IV hemagglutinin for α2,3 versus α2,6 sialic acid linkage to galactose; flavivirus mutants with altered neurovirulence	Several studies reviewed in Reference 1
Substitutions in B cell or T cell epitopes Antigenic variation (drift)	Evasion of antibody response; most studies with IV but described for many viruses; CTL escape; antigenic variation may be linked to changes in cell tropism	Several studies reviewed in Reference 1

Only a few of many examples of relevant effects of single mutations or amino acid substitutions reported in the literature are included here. Abbreviations: CTL, cytotoxic T cell; HIV-1, human immunodeficiency virus type 1; IV, influenza virus; SIV, simian immunodeficiency virus.

IFN-induction capacity have been retrieved from viral populations and the relevant mutations introduced into live-attenuated vaccine candidates (51). Mutant spectra can be a source of interesting and useful virus variants.

From the architecture of mutant swarms, it follows that new traits whose acquisition depends on a single mutation are particularly prone to be gained or lost upon shifts in genome dominance. Phenotypic modifications due to single mutations in viral genomes are continuously being reported; some examples are listed in **Table 1**. The relevant amino acid replacements may be conservative or drastic, and they may affect conserved or variable protein domains; no rules have become apparent. Dependence on single mutations concomitantly with protein multifunctionality facilitates mutual influences among traits. A documented case is coevolution of cell tropism and antigenicity, which is aided by the overlap between antigenic sites and receptor recognition domains on surface virion proteins (52). Therefore, the mutant spectrum composition may simultaneously participate in the evasion and the modulatory strategies of virus interaction with the immune system.

6. QUASISPECIES IN VIRUS-HOST INTERACTIONS

Penetration of viruses into their host cells triggers a cascade of mutual virus-cell influences, reflected in modifications of the cell expression program and its derived metabolic pathways. The virus in turn tends to accommodate its mutant spectrum to the cellular environment and then to

accessible and susceptible host compartments. The effects on the host can lead to post-infection sequels and pathologies, so that the consequences of an infection can last even after the virus is no longer detected (even perhaps after complete elimination).

For example, the cellular perturbations associated with hepatitis C virus (HCV) include induction of inflammatory and oxidative stress responses, host DNA and mitochondrial damage, aberrant expression of microRNAs, and increased hepatocyte proliferation (53). A high HCV quasispecies complexity has been connected with liver disease progression and poor response to treatment (54–56). Epigenetic cell modifications can enhance cancer risk after a sustained virological response (57). In some analyses of hepatocellular carcinomas (HCCs) developed during active viral replication, the HCV quasispecies complexity was larger within the tumor than in surrounding tissue (58), suggesting that the virus may respond to an unfavorable tumor environment by genome diversification. Hepatitis B virus (HBV) quasispecies complexity has also been related to progression toward HCC (59), and the quasispecies complexity of the HBV X gene (HBX) integrated in the tumor DNA of cancer patients was larger than the complexity of unintegrated HBX (60). Immunodeficiency-related vaccine-derived rubella virus in children displays high rates of intrahost evolution, and the virus in cutaneous granulomas shows a broader mutant spectrum (61).

Studies with plant viruses have documented that the patterns of host perturbations, including epigenetic modifications, are influenced by the genotype profile of the virus and its degree of adaptation to the host (62–65).

Several mechanisms may be responsible for virus-induced host modifications: viral fitness per se (measured as the rate of intracellular replication), the intracellular viral load (which is related to fitness), the presence of specific variants, combinations of variants, or modulation by defective genomes (66), among other influences. These possibilities illustrate the challenges researchers confront when trying to understand the interactions of mutant swarms with cells.

Regarding adjustment of mutant spectra to environmental demands, several analyses have documented differences in mutant spectrum composition at different sites of the same infected host individual. Among many examples, an early study documented selection of organ-specific lymphocytic choriomeningitis virus variants (67). Distinct HIV-1 mutant spectra were located in different anatomic compartments or human brain regions (68, 69), and in tissue-specific quasispecies, syncytium-inducing and noninducing HIV-1 phenotypes coexisted (70); mutant spectrum complexity was associated with progression toward acquired immunodeficiency syndrome (AIDS) and neuropsychological impairment (71). HIV-1 compartmentalization may be favored by differential coreceptor expression. The size of the West Nile virus (WNV) mutant spectrum correlated with virus ability to replicate in mosquito cells (72). A distinct mutant composition in different compartments may have its origin in the colonization of a new site following a bottleneck known to occur within infected organisms (73, 74) and not only in a direct modification of an entire mutant swarm in the new environment. This duality is a reflection of the old neutralist-selectionist controversy.

Despite the flexibility conferred by a broad mutant spectrum, it is not a predictor of general adaptability per se. The size of the WNV mutant spectrum attained in *Culex pipiens* or young chicken hosts correlated negatively with mortality in mice (75). Adaptation of mutant spectra to alternative environments is a mechanism of virus attenuation that operates in the preparation of live-attenuated vaccines that can be regarded as reshaped mutant swarms; virulent and attenuated viruses may coexist in vaccine preparations (76, 77).

With a mouse-adapted foot-and-mouth disease virus, mutant spectra displaying distinctive virulence were characterized in different mouse organs (78). The subpopulations were indistinguishable by their consensus sequences, and had similar mutant spectrum complexities, within the resolution of the assay. The results mean that virulence determinants may be absent or present in

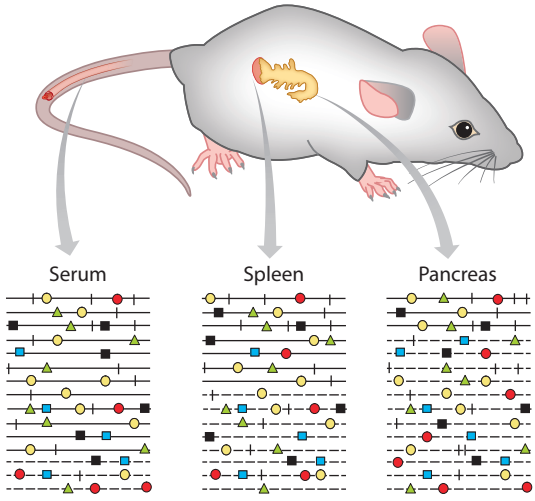

Figure 3

A simplified representation of compartmentalized quasispecies in a mouse infected with foot-and-mouth disease virus (FMDV) (78). Virus pathogenicity can be the result of complex interactions among components of the mutant spectrum. In the mutant distributions (depicted as in **Figure 1*b***), solid lines represent pathogenic genomes and dashed lines represent nonpathogenic genomes. The model is intended to explain different virulence for mice of three FMDV populations that display equal consensus sequence and population complexity. Figure adapted with permission from Reference 78.

mutant spectra of comparable complexity (**Figure 3**). Viral virulence may be dictated not only by specific viral and host genes (as often assumed) but also by population parameters and the mutant spectrum composition.

7. A PERMANENT POPULATION DISEQUILIBRIUM

Several recent studies have addressed quasispecies dynamics by deep sequencing analysis of sequential viral populations replicating in cell culture (43, 79, 80). Replication in established cell lines is as close as presently can be devised for a virus to approach some equilibrium with an environment. Yet, the dynamics observed is of relentless mutant spectrum modifications. Mutant spectrum analysis of HCV replicating in human hepatoma cells in culture indicated that after 200 serial infections (which are equivalent to about 700 days of uninterrupted viral replication), changes in mutant frequency of individual residues were as prominent as in the initial passages, if not more, and the modifications brought about phenotypic changes (81). This study suggests that RNA viruses, whether replicating in changing or static environments, keep a sustained level of genetic heterogeneity that maintains their capacity to respond to constraints, independent of a history of environmental invariance. If transient equilibria exist (Section 2), they are of very short duration relative to the time frame of laboratory observations. Procedures to dissect underlying groups of related sequences within vastly complex mutant spectra have been developed (82–85).

A state of population disequilibrium weakens the value of a master or dominant genome as a guarantee of population resilience against increases of mutation rate. The HCV data suggest two possible influences on the stability of the genetic information conveyed by viral populations (**Figure 2**; Sections 2 and 9): (*a*) the superiority of the master (dominant) genome and (*b*) the accommodation of the entire mutant spectrum to an environment, even if successive master sequences have only a transient existence. Most likely, factor (*b*) is the major one involved

in fitness gain and resistance to lethal mutagenesis (Section 9). The reason is that long-term adaptation of HCV to a cell culture environment was associated with an expansion of the mutant spectrum rather than a convergence toward a single new consensus sequence (79, 81). Again, mutant spectra play a prominent role.

8. THE ELUSIVE MEANING OF CONSERVATION

The multiple low-frequency variants found in mutant spectra point toward a need to reconsider the concept of lethality. With the resolution power of deep sequencing, some mutations that were previously cataloged as lethal may instead be highly deleterious (fitness decreasing) but not lethal. This puzzle has also been raised for bacterial populations (86). A related aspect refers to which nucleotide or amino acid sequences can be cataloged as conserved (invariant) among members of a given virus. Residues that are considered conserved according to the alignment of genomic sequences in data banks turn out to be variable when the alignment is done with sequences retrieved from mutant spectra. This difference has been documented with HCV replicating in cell culture or in infected patients (87). The study involved a two-rank comparison of conservation scores: between mutant spectra and their corresponding consensus sequences, and between mutant spectra and Los Alamos National Laboratory HCV sequence alignment. Residue conservation according to consensus sequences or data banks does not necessarily correspond to conservation at the quasispecies level. A relaxed negative selection acting at the quasispecies level versus multiple selective and bottleneck filters undergone by viral populations until a genomic sequence is incorporated in data banks may provide an explanation. The implications are far from trivial. The standard calculation of nucleotide or amino acid conservation for the design of broad spectrum (or universal) vaccines or antiviral agents may not diminish the probability of selection of escape mutants. The putative universal antiviral reagents will not be directed to consensus sequences but to mutant spectra replicating in infected individuals. Despite their likely inferior replicative fitness, minority genomes that are not represented in data banks may be only a short mutational distance away from gaining fitness, thereby becoming competent escape mutants. Design of universal ligands remains an important challenge, and, if undertaken, it should be endorsed by a more stringent conservation evaluation using mutant spectra data in the sequence alignment (87). If the results with HCV are extended to other viruses, they will diminish the scope of residue conservation of viruses to a far more reduced subset than presently thought. The observation adds to the limited value of using a consensus sequence as the representative of a viral population (33, 88–90).

9. IMPACT ON ANTIVIRAL STRATEGIES

Successful vaccines and antiviral treatments have been developed for decades, without awareness of mutant swarms and their implications. Yet, at least part of their failures and limitations has been attributed to not having calibrated sufficiently the genome variation dynamics of viruses. Benefits will follow from considering viruses as true moving targets and from refraining from combating complexity with simplicity (monotherapy, simple vaccine formulations) (1, 33). Caution has also been expressed with designs intended to control the COVID-19 pandemic (91; see also Section 10).

In this scenario, lethal mutagenesis or the extinction of viruses by an excess of mutations driven by nucleotide analogs is gradually expanding as a broad-spectrum strategy applicable to emerging viral infections (**Figure 1**; Section 2). The first results of an adverse effect of increased mutation rates on virus viability were obtained by J.J. Holland and associates (92), and the term lethal mutagenesis was coined by Lawrence Loeb and colleagues (93), who also performed the

first clinical trial with AIDS patients (94) and opened the approach for anticancer therapy (95). The virus transition toward extinction is a complex multistep process, initiated by defective genomes and culminating in overt lethality. The main features of lethal mutagenesis of viruses are reviewed in Reference 96 and include the following:

- It is an intracellular event that is induced by increases of mutation rate, often by nucleotide analogs.
- It requires mutagenic activity. Equivalent inhibitory levels without mutagenesis are not effective. It is not due to a mere R_0 reduction.
- It is characterized by an increase of mutation frequency and a 10^2- to 10^3-fold decrease of specific infectivity without modification of the consensus sequence.
- It is a multistage process involving at least two recognized steps: (*a*) lethal defection (generation of defective mutants that interfere with replication of the ensemble) and (*b*) overt lethality (number of mutations incompatible with completion of the infectious cycle).
- It has proven effective against several RNA viruses in cell culture and in vivo, including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Some antivirals that are licensed for human use and that have been used in antiviral treatments (favipiravir, ribavirin) may have been effective partly due to lethal mutagenesis.
- Synergisms (among lethal mutagens, and between lethal mutagens and nonmutagenic inhibitors) offer additional promise as broad spectrum treatments against emerging viral pathogens.
- Several compatible mechanisms based on quasispecies theory and the population genetics concept have been proposed to explain the experimental and clinical observations (reviewed in Reference 8).

These features, along with the accumulated evidence in cell culture and in vivo, render this approach a promising one to be added to the traditional combination therapies with nonmutagenic inhibitors.

10. ANY LESSONS FOR COVID-19?

When the possible connections between quasispecies dynamics and long-term evolution at the epidemiological level had been explicitly recognized as a pertinent question (87, 97–99), COVID-19 erupted in the human population. This new pandemic associated with the betacoronavirus SARS-CoV-2 has posed the question of to what extent the 3′ to 5′ exonuclease activity encoded by the SARS-CoV-2 genome (domain present in protein nsp14) decreases the mutation rate inherent to the replicative complex (that includes nsp12 and several cofactors). In other coronaviruses, the equivalent 3′ to 5′ exonuclease activity lowers the mutation rate about 15-fold, and such a decrease is thought to contribute to the genetic stability of the large coronavirus RNA genome (100). At the time of this writing, to our knowledge, the mutation rate of SARS-CoV-2 in absence and presence of an active 3′ to 5′ exonuclease has not been calculated. The mutation rate for SARS-CoV-2 and its consequences for mutant spectrum broadness are relevant to a virus with a large RNA genome that can colonize different human organs (101).

Perhaps due to the urgency to understand basic replicative features of the virus and to respond to the public health emergency, expressions of concerns about the impact of genetic and antigenic diversity for antiviral and vaccine efficacy have been limited, despite decades of experience with influenza vaccines. Also, there is evidence of S (spike protein) diversification through point mutations and insertions-deletions, with amino acid substitutions that may confer resistance to antibodies and convalescent plasma (102). SARS-CoV-2 replicates as mutant spectra (103, 104),

as previously reported for the Middle East respiratory syndrome coronavirus (105, 106). With a mutant frequency cutoff of 1%, mutant spectrum differences were revealed among viruses from different anatomical compartments and among sequential virus samples from the same patient (107, 108).

Comparison of genomic consensus sequences (that gradually replenish the GISAID and other databases) yields an average rate of SARS-CoV-2 evolution at the epidemiological (interhost) level of $(1.2 \pm 0.5) \times 10^{-3}$ mutations per site and year (range 9.9×10^{-4} to 2.2×10^{-3} mutations per site and year) [this calculation is based on data from 10 reports (L. Vázquez-Sirvent, B. Martínez-González, M.E. Soria, C. Perales, unpublished results)]. The rate of evolution and the estimates of intrahost heterogeneity indicate that even a large viral RNA genome follows the typical evolutionary patterns described for other RNA viruses. The possibility that SARS-CoV-2 replicates with an average error rate close to the maximum acceptable to maintain functionality is suggested by its vulnerability to lethal mutagenesis when the full catalytic potential is expressed by the nsp12-nsp7-nsp8 polymerase complex (17, 109). The analog favipiravir—which exerts its antiviral activity at least partly by lethal mutagenesis—has proven effective in some clinical assays (110, 111), therefore adding to treatment options for COVID-19 that may also be of value for future viral emergences.

11. CLOSING STATEMENTS

The theoretical and experimental origins of viral quasispecies are a demonstration of the benefits that can be gained by bridging research fields. The convergence between a model of the origin of life and the empirical observations on the mutant spectrum nature of viral populations has unveiled a huge layer of minority genomes that were hidden from inspection until clonal analyses and deep sequencing were used. Covert mutant swarms can either modulate the behavior of the population ensemble or become prominent as a result of selection or random events such as bottleneck isolations of individual variants.

It is expected that increasingly powerful and accurate sequencing methods, combined with analyses of viral genomes in single infected cells, will provide a realistic picture of the nature of the viral populations we are confronting. We should worry about individual mutations (despite being unable to prevent them) because what is minority at a given moment may be prominent the next. Consensus sequences provide an identification but not a characterization of a virus. There is a clear trend toward planning antiviral strategies according to the challenge posed by quasispecies dynamics. It is hoped that a more detailed description of mutant spectra, in parallel with the accommodation of vaccine and treatment strategies to the reality of viral populations, will improve both our knowledge of viruses and our capacity to react to disease challenges.

SUMMARY POINTS

1. Viral quasispecies are the dynamic mutant distributions that replicate in infected hosts. The concept had a theoretical origin in a model of the origin of life and an independent origin in observations on the presence of multiple variant genomes in RNA bacteriophage populations.
2. The major impact of quasispecies for virology lies in the consideration of the wild type as a set of related genomes, the interpretation of virus adaptability by modification of the composition of mutant spectra that is accelerated by high mutation rates, and new possibilities of antiviral interventions.

3. Mutant spectra are reservoirs of phenotypically relevant variants that may affect the interaction with the host, disease processes, and response to treatments. Mutant genomes may act collectively as units of selection.
4. A consensus sequence is a simplified representation of a viral population. It may not be present in the virus population it intends to represent. Studies with hepatitis C virus suggest that residue conservation according to alignments of consensus sequences or sequences compiled in data banks does not fit conservation of the same residues in mutant spectra.
5. Lethal mutagenesis is an antiviral design inspired in the error catastrophe concept of quasispecies theory. It consists of the extinction of a virus by introducing in the replicating viral genome a number of mutations incompatible with maintenance of infectivity. There is increasing evidence of its effectiveness in cell culture and in vivo.

FUTURE ISSUES

1. Clarification is needed of the extent of operation and biological implications of quasispecies dynamics for complex DNA viruses and cellular populations, in particular tumor cells.
2. Application of ultra-deep sequencing techniques and new bioinformatics approaches to achieve deeper penetration into the composition of mutant spectra is needed, in particular the use of whole-genome sequencing to identify mutation linkage and to interpret fitness effects of epistasis in dominant and minority genomes.
3. Implementation of expanded data banks for viral genomes that incorporate mutant spectrum sequences for a stricter criterion of residue conservation should help in the design of universal ligands and vaccines.
4. It will be necessary to distinguish viable and defective genomes in the mutant spectra characterized by deep sequencing and the role of defective genomes in the interactions of the viable virus subset with its host.
5. Advances in lethal mutagenesis designs should center on finding (or designing) new mutagenic agents that increase the error rate of the viral (but not of the cellular) polymerases and that can act synergistically to avoid selection of escape viral mutants.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We are indebted to many students and colleagues for work on the quasispecies concept. The research at CBMSO was supported by grants SAF2014-52400-R from Ministerio de Economía y Competitividad (MINECO), SAF2017-87846-R and BFU2017-91384-EXP from Ministerio

de Ciencia, Innovación y Universidades (MCIU), PI18/00210 from Instituto de Salud Carlos III, S2013/ABI-2906 (PLATESA from Comunidad de Madrid/FEDER), and S2018/BAA-4370 (PLATESA2 from Comunidad de Madrid/FEDER). C.P. is supported by the Miguel Servet program of the Instituto de Salud Carlos III (CP14/00121 and CPII19/00001), cofinanced by the European Regional Development Fund (ERDF). CIBERehd (Centro de Investigación en Red de Enfermedades Hepáticas y Digestivas) is funded by Instituto de Salud Carlos III. Institutional grants from the Fundación Ramón Areces and Banco Santander to CBMSO are also acknowledged. The team at CBMSO belongs to the Global Virus Network (GVN). C.G.-C. is supported by predoctoral contract PRE2018-083422 from MCIU.

LITERATURE CITED

1. Domingo E. 2020. *Virus as Populations*. Amsterdam: Elsevier. 2nd ed.
2. Holland JJ, Spindler K, Horodyski F, Grabau E, Nichol S, VandePol S. 1982. Rapid evolution of RNA genomes. *Science* 215:1577–85
3. Eigen M. 1971. Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58:465–523
4. Eigen M, Schuster P. 1979. *The Hypercycle: A Principle of Natural Self-Organization*. Berlin: Springer
5. Mills DR, Peterson RL, Spiegelman S. 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule. *PNAS* 58:217–24
6. Bull JJ, Meyers LA, Lachmann M. 2005. Quasispecies made simple. *PLOS Comput. Biol.* 1:e61
7. Tejero H, Montero F, Nuno JC. 2016. Theories of lethal mutagenesis: from error catastrophe to lethal defection. *Curr. Top. Microbiol. Immunol.* 392:161–79
8. Domingo E, Schuster P. 2016. *Quasispecies: From Theory to Experimental Systems*. Cham, Switz.: Springer
9. Szostak N, Wasik S, Blazewicz J. 2016. Hypercycle. *PLOS Comput. Biol.* 12:e1004853
10. Eigen M. 2002. Error catastrophe and antiviral strategy. *PNAS* 99:13374–76
11. Eigen M. 2013. *From Strange Simplicity to Complex Familiarity*. Oxford, UK: Oxford Univ. Press
12. Weissmann C, Billeter MA, Goodman HM, Hindley J, Weber H. 1973. Structure and function of phage RNA. *Annu. Rev. Biochem.* 42:303–28
13. Weissmann C. 1974. The making of a phage. *FEBS Lett.* 40(S1):S3–9
14. Takeshita D, Yamashita S, Tomita K. 2014. Molecular insights into replication initiation by Qβ replicase using ribosomal protein S1. *Nucleic Acids Res.* 42:10809–22
15. Flavell RA, Sabo DL, Bandle EF, Weissmann C. 1974. Site-directed mutagenesis: generation of an extracistronic mutation in bacteriophage Qβ RNA. *J. Mol. Biol.* 89:255–72
16. Domingo E, Flavell RA, Weissmann C. 1976. In vitro site-directed mutagenesis: generation and properties of an infectious extracistronic mutant of bacteriophage Qβ. *Gene* 1:3–25
17. Sheahan TP, Sims AC, Zhou S, Graham RL, Pruijssers AJ, et al. 2020. An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice. *Sci. Transl. Med.* 12:eabb5883
18. Agostini ML, Pruijssers AJ, Chappell JD, Gribble J, Lu X, et al. 2019. Small-molecule antiviral β-D-N^4-hydroxycytidine inhibits a proofreading-intact coronavirus with a high genetic barrier to resistance. *J. Virol.* 93:e01348-19
19. de Wachter R, Fiers W. 1972. Preparative two-dimensional polyacrylamide gel electrophoresis of ^{32}P-labeled RNA. *Anal. Biochem.* 49:184–97
20. Billeter MA. 1978. Sequence and location of large RNase T1 oligonucleotides in bacteriophage Qbeta RNA. *J. Biol. Chem.* 253:8381–89
21. Weissmann C, Tanaguchi T, Domingo E, Sabo D, Flavell RA. 1977. Site-directed mutagenesis as a tool in genetics. In *Genetic Manipulation as It Affects the Cancer Problem*, ed. J Schultz, Z Brada, pp. 11–36. New York: Academic
22. Batschelet E, Domingo E, Weissmann C. 1976. The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. *Gene* 1:27–32

23. Drake JW. 1969. Comparative rates of spontaneous mutation. *Nature* 221:1132
24. Drake JW, Charlesworth B, Charlesworth D, Crow JF. 1998. Rates of spontaneous mutation. *Genetics* 148:1667–86
25. Drake JW, Holland JJ. 1999. Mutation rates among RNA viruses. *PNAS* 96:13910–13
26. Sanjuan R, Nebot MR, Chirico N, Mansky LM, Belshaw R. 2010. Viral mutation rates. *J. Virol.* 84:9733–48
27. Domingo E, Sabo D, Taniguchi T, Weissmann C. 1978. Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13:735–44
28. Biebricher CK, Eigen M, Gardiner WC Jr. 1985. Kinetics of RNA replication: competition and selection among self-replicating RNA species. *Biochemistry* 24:6550–60
29. Domingo E, Holland JJ, Ahlquist P. 1988. *RNA Genetics*. Boca Raton, FL: CRC
30. Holmes EC, Moya A. 2002. Is the quasispecies concept relevant to RNA viruses? *J. Virol.* 76(1):460–62
31. Domingo E. 2002. Quasispecies theory in virology. *J. Virol.* 76: 463–65
32. de la Torre JC, Holland JJ. 1990. RNA virus quasispecies populations can suppress vastly superior mutant progeny. *J. Virol.* 64:6278–81
33. Domingo E, Perales C. 2019. Viral quasispecies. *PLOS Genet*. 15:e1008271
34. Domingo E, Schuster P. 2016. What is a quasispecies? Historical origins and current scope. *Curr. Top. Microbiol. Immunol.* 392:1–22
35. Briones C, Domingo E. 2008. Minority report: hidden memory genomes in HIV-1 quasispecies and possible clinical implications. *AIDS Rev*. 10:93–109
36. Sukkaew A, Thanagith M, Thongsakulprasert T, Mutso M, Mahalingam S, et al. 2018. Heterogeneity of clinical isolates of chikungunya virus and its impact on the responses of primary human fibroblast-like synoviocytes. *J. Gen. Virol.* 99:525–35
37. Yin J. 1993. Evolution of bacteriophage T7 in a growing plaque. *J. Bacteriol.* 175:1272–77
38. Voorhees IEH, Lee H, Allison AB, Lopez-Astacio R, Goodman LB, et al. 2019. Limited intrahost diversity and background evolution accompany 40 years of canine parvovirus host adaptation and spread. *J. Virol.* 94:e01162-19
39. Jordan-Paiz A, Nevot M, Lamkiewicz K, Lataretu M, Franco S, et al. 2020. HIV-1 lethality and loss of Env protein expression induced by single synonymous substitutions in the virus genome intronic-splicing silencer. *J. Virol.* 94:e01108-20
40. Nam B, Mekuria Z, Carossino M, Li G, Zheng Y, et al. 2019. Intrahost selection pressure drives equine arteritis virus evolution during persistent infection in the stallion reproductive tract. *J. Virol.* 93:e00045-19
41. Honce R, Schultz-Cherry S. 2020. They are what you eat: shaping of viral populations through nutrition and consequences for virulence. *PLOS Pathog.* 16:e1008711
42. Mira A, Martin-Cuadrado AB, D'Auria G, Rodriguez-Valera F. 2010. The bacterial pan-genome: a new paradigm in microbiology. *Int. Microbiol.* 13:45–57
43. Acevedo A, Brodsky L, Andino R. 2014. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* 505:686–90
44. Riaz N, Leung P, Barton K, Smith MA, Carswell S, et al. 2021. Adaptation of Oxford Nanopore technology for hepatitis C whole genome sequencing and identification of within-host viral variants. *BMC Genom.* 22(1):148
45. Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, et al. 1997. Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nat. Med.* 3:205–11
46. Young DF, Wignall-Fleming EB, Busse DC, Pickin MJ, Hankinson J, et al. 2019. The switch between acute and persistent paramyxovirus infection caused by single amino acid substitutions in the RNA polymerase P subunit. *PLOS Pathog.* 15:e1007561
47. Oldstone MB. 2006. Viral persistence: parameters, mechanisms and future predictions. *Virology* 344:111–18
48. Marcus PI, Rodriguez LL, Sekellick MJ. 1998. Interferon induction as a quasispecies marker of vesicular stomatitis virus populations. *J. Virol.* 72:542–49

49. Garcin D, Itoh M, Kolakofsky D. 1997. A point mutation in the Sendai virus accessory C proteins attenuates virulence for mice, but not virus growth in cell culture. *Virology* 238:424–31
50. Young DF, Chatziandreou N, He B, Goodbourn S, Lamb RA, Randall RE. 2001. Single amino acid substitution in the V protein of simian virus 5 differentiates its ability to block interferon signaling in human and murine cells. *J. Virol.* 75:3363–70
51. Ghorbani A, Abundo MC, Ji H, Taylor KJM, Ngunjiri JM, Lee CW. 2020. Viral subpopulation screening guides in designing a high interferon inducing live attenuated influenza vaccine by targeting rare mutations in NS1 and PB2 proteins. *J. Virol.* 95(2):e01722-20
52. Baranowski E, Ruiz-Jarabo CM, Pariente N, Verdaguer N, Domingo E. 2003. Evolution of cell recognition by viruses: a source of biological novelty with medical implications. *Adv. Virus Res.* 62:19–111
53. Ringelhan M, McKeating JA, Protzer U. 2017. Viral hepatitis and liver cancer. *Philos. Trans. R. Soc. B* 372:20160274
54. Koizumi K, Enomoto N, Kurosaki M, Murakami T, Izumi N, et al. 1995. Diversity of quasispecies in various disease stages of chronic hepatitis C virus infection and its significance in interferon treatment. *Hepatology* 22:30–35
55. Farci P, Shimoda A, Coiana A, Diaz G, Peddis G, et al. 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science* 288:339–44
56. Kumar D, Malik A, Asim M, Chakravarti A, Das RH, Kar P. 2008. Influence of quasispecies on virological responses and disease severity in patients with chronic hepatitis C. *World J. Gastroenterol.* 14:701–8
57. Hamdane N, Juhling F, Crouchet E, El Saghire H, Thumann C, et al. 2019. HCV-induced epigenetic changes associated with liver cancer risk persist after sustained virologic response. *Gastroenterology* 156:2313–29.e7
58. Horie C, Iwahana H, Horie T, Shimizu I, Yoshimoto K, et al. 1996. Detection of different quasispecies of hepatitis C virus core region in cancerous and noncancerous lesions. *Biochem. Biophys. Res. Commun.* 218:674–81
59. Zhang AY, Lai CL, Huang FY, Seto WK, Fung J, et al. 2017. Deep sequencing analysis of quasispecies in the HBV pre-S region and its association with hepatocellular carcinoma. *J. Gastroenterol.* 52:1064–74
60. Fang X, Wu HH, Ren JJ, Liu HZ, Li KZ, et al. 2017. Associations between serum HBX quasispecies and their integration in hepatocellular carcinoma. *Int. J. Clin. Exp. Pathol.* 10:11857–66
61. Perelygina L, Chen MH, Suppiah S, Adebayo A, Abernathy E, et al. 2019. Infectious vaccine-derived rubella viruses emerge, persist, and evolve in cutaneous granulomas of children with primary immunodeficiencies. *PLOS Pathog.* 15:e1008080
62. Agudelo-Romero P, Carbonell P, de la Iglesia F, Carrera J, Rodrigo G, et al. 2008. Changes in the gene expression profile of *Arabidopsis thaliana* after infection with *Tobacco etch virus*. *Virol. J.* 5:92
63. Hillung J, Garcia-Garcia F, Dopazo J, Cuevas JM, Elena SF. 2016. The transcriptomics of an experimentally evolved plant-virus interaction. *Sci. Rep.* 6:24901
64. Cervera H, Ambros S, Bernet GP, Rodrigo G, Elena SF. 2018. Viral fitness correlates with the magnitude and direction of the perturbation induced in the host's transcriptome: the tobacco etch potyvirus—tobacco case study. *Mol. Biol. Evol.* 35:1599–615
65. Correa RL, Sanz-Carbonell A, Kogej Z, Muller SY, Ambros S, et al. 2020. Viral fitness determines the magnitude of transcriptomic and epigenomic reprograming of defense responses in plants. *Mol. Biol. Evol.* 37:1866–81
66. Vignuzzi M, Lopez CB. 2019. Defective viral genomes are key drivers of the virus-host interaction. *Nat. Microbiol.* 4:1075–87
67. Ahmed R, Oldstone MB. 1988. Organ-specific selection of viral variants during chronic infection. *J. Exp. Med.* 167:1719–24
68. Wong JK, Ignacio CC, Torriani F, Havlir D, Fitch NJ, Richman DD. 1997. In vivo compartmentalization of human immunodeficiency virus: evidence from the examination of pol sequences from autopsy tissues. *J. Virol.* 71:2059–71
69. Shapshak P, Segal DM, Crandall KA, Fujimura RK, Zhang BT, et al. 1999. Independent evolution of HIV type 1 in different brain regions. *AIDS Res. Hum. Retroviruses* 15:811–20

70. van't Wout AB, Ran LJ, Kuiken CL, Kootstra NA, Pals ST, Schuitemaker H. 1998. Analysis of the temporal relationship between human immunodeficiency virus type 1 quasispecies in sequential blood samples and various organs obtained at autopsy. *J. Virol.* 72:488–96
71. Hightower GK, Wong JK, Letendre SL, Umlauf AA, Ellis RJ, et al. 2012. Higher HIV-1 genetic diversity is associated with AIDS and neuropsychological impairment. *Virology* 433:498–505
72. Ciota AT, Ngo KA, Lovelace AO, Payne AF, Zhou Y, et al. 2007. Role of the mutant spectrum in adaptation and replication of West Nile virus. *J. Gen. Virol.* 88:865–74
73. Pfeiffer JK, Kirkegaard K. 2006. Bottleneck-mediated quasispecies restriction during spread of an RNA virus from inoculation site to brain. *PNAS* 103:5520–25
74. Bull RA, Luciani F, McElroy K, Gaudieri S, Pham ST, et al. 2011. Sequential bottlenecks drive viral evolution in early acute hepatitis C virus infection. *PLOS Pathog.* 7:e1002243
75. Jerzak GV, Bernard K, Kramer LD, Shi PY, Ebel GD. 2007. The West Nile virus mutant spectrum is host-dependant and a determinant of mortality in mice. *Virology* 360:469–76
76. Chumakov KM, Powers LB, Noonan KE, Roninson IB, Levenbook IS. 1991. Correlation between amount of virus with altered nucleotide sequence and the monkey test for acceptability of oral poliovirus vaccine. *PNAS* 88:199–203
77. Marcus PI, Ngunjiri JM, Sekellick MJ, Wang L, Lee CW. 2010. *In vitro* analysis of virus particle subpopulations in candidate live-attenuated influenza vaccines distinguishes effective from ineffective vaccines. *J. Virol.* 84:10974–81
78. Sanz-Ramos M, Diaz-San Segundo F, Escarmis C, Domingo E, Sevilla N. 2008. Hidden virulence determinants in a viral quasispecies in vivo. *J. Virol.* 82:10465–76
79. Moreno E, Gallego I, Gregori J, Lucia-Sanz A, Soria ME, et al. 2017. Internal disequilibria and phenotypic diversification during replication of hepatitis C virus in a noncoevolving cellular environment. *J. Virol.* 91:e02505-16
80. Braun T, Borderia AV, Barbezange C, Vignuzzi M, Louzoun Y. 2019. Long-term context-dependent genetic adaptation of the viral genetic cloud. *Bioinformatics* 35:1907–15
81. Gallego I, Soria ME, Garcia-Crespo C, Chen Q, Martinez-Barragan P, et al. 2020. Broad and dynamic diversification of infectious hepatitis C virus in a cell culture environment. *J. Virol.* 94:e01856-19
82. Baccam P, Thompson RJ, Fedrigo O, Carpenter S, Cornette JL. 2001. PAQ: partition analysis of quasispecies. *Bioinformatics* 17:16–22
83. Skums P, Zelikovsky A, Singh R, Gussler W, Dimitrova Z, et al. 2018. QUENTIN: reconstruction of disease transmissions from viral quasispecies genomic data. *Bioinformatics* 34:163–70
84. Ahn S, Ke Z, Vikalo H. 2018. Viral quasispecies reconstruction via tensor factorization with successive read removal. *Bioinformatics* 34:i23–31
85. Henningsson R, Moratorio G, Borderia AV, Vignuzzi M, Fontes M. 2019. DISSEQT—DIStribution-based modeling of SEQuence space Time dynamics. *Virus Evol.* 5:vez028
86. Jee J, Rasouly A, Shamovsky I, Akivis Y, Steinman SR, et al. 2016. Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing. *Nature* 534:693–96
87. Garcia-Crespo C, Soria ME, Gallego I, Avila AI, Martinez-Gonzalez B, et al. 2020. Dissimilar conservation pattern in hepatitis C virus mutant spectra, consensus sequences, and data banks. *J. Clin. Med.* 9:3450
88. Domingo E. 2006. Quasispecies: concepts and implications for virology. *Curr. Top. Microbiol. Immunol.* 299:1–31
89. Caldwell HS, Lasek-Nesselquist E, Follano P, Kramer LD, Ciota AT. 2020. Divergent mutational landscapes of consensus and minority genotypes of West Nile virus demonstrate host and gene-specific evolutionary pressures. *Genes* 11:1299
90. Knyazev S, Hughes L, Skums P, Zelikovsky A. 2020. Epidemiological data analysis of viral quasispecies in the next-generation sequencing era. *Briefings Bioinform.* 22:96–108
91. Kennedy DA, Read AF. 2020. Monitor for COVID-19 vaccine resistance evolution during clinical trials. *PLOS Biol.* 18:e3001000
92. Holland JJ, Domingo E, de la Torre JC, Steinhauer DA. 1990. Mutation frequencies at defined single codón sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J. Virol.* 64:3960–62

93. Loeb LA, Essigmann JM, Kazazi F, Zhang J, Rose KD, Mullins JI. 1999. Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *PNAS* 96:1492–97
94. Mullins JI, Heath L, Hughes JP, Kicha J, Styrchak S, et al. 2011. Mutation of HIV-1 genomes in a clinical population treated with the mutagenic nucleoside KP1461. *PLOS ONE* 6:e15135
95. Loeb LA. 2011. Human cancers express mutator phenotypes: origin, consequences and targeting. *Nat. Rev. Cancer* 11:450–57
96. Perales C, Gallego I, de Avila AI, Soria ME, Gregori J, et al. 2019. The increasing impact of lethal mutagenesis of viruses. *Future Med. Chem.* 11:1645–57
97. Geoghegan JL, Holmes EC. 2018. Evolutionary virology at 40. *Genetics* 210:1151–62
98. Domingo E, Soria ME, Gallego I, de Avila AI, Garcia-Crespo C, et al. 2020. A new implication of quasispecies dynamics: broad virus diversification in absence of external perturbations. *Infect. Genet. Evol.* 82:104278
99. Glebova O, Knyazev S, Melnyk A, Artyomenko A, Khudyakov Y, et al. 2017. Inference of genetic relatedness between viral quasispecies from sequencing data. *BMC Genom.* 18:918
100. Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS. 2011. Coronaviruses: An RNA proofreading machine regulates replication fidelity and diversity. *RNA Biol.* 8(2):270–79
101. Trypsteen W, Van Cleemput J, Snippenberg WV, Gerlo S, Vandekerckhove L. 2020. On the whereabouts of SARS-CoV-2 in the human body: a systematic review. *PLOS Pathog.* 16:e1009037
102. Weisblum Y, Schmidt F, Zhang F, DaSilva J, Poston D, et al. 2020. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* 9:e61312
103. Karamitros T, Papadopoulou G, Bousali M, Mexias A, Tsiodras S, Mentis A. 2020. SARS-CoV-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies. *J. Clin. Virol.* 131:104585
104. Capobianchi MR, Rueca M, Messina F, Giombini E, Carletti F, et al. 2020. Molecular characterization of SARS-CoV-2 from the first case of COVID-19 in Italy. *Clin. Microbiol. Infect.* 26:954–56
105. Xu D, Zhang Z, Wang FS. 2004. SARS-associated coronavirus quasispecies in individual patients. *N. Engl. J. Med.* 350:1366–67
106. Park D, Huh HJ, Kim YJ, Son DS, Jeon HJ, et al. 2016. Analysis of intrapatient heterogeneity uncovers the microevolution of Middle East respiratory syndrome coronavirus. *Cold Spring Harb. Mol. Case Stud.* 2:a001214
107. Jary A, Leducq V, Malet I, Marot S, Klement-Frutos E, et al. 2020. Evolution of viral quasispecies during SARS-CoV-2 infection. *Clin. Microbiol. Infect.* 26:1560.e1–4
108. Rueca M, Bartolini B, Gruber CEM, Piralla A, Baldanti F, et al. 2020. Compartmentalized replication of SARS-Cov-2 in upper versus lower respiratory tract assessed by whole genome quasispecies analysis. *Microorganisms* 8:1302
109. Shannon A, Selisko B, Le NT, Huchting J, Touret F, et al. 2020. Rapid incorporation of favipiravir by the fast and permissive viral RNA polymerase complex results in SARS-CoV-2 lethal mutagenesis. *Nat. Commun.* 11:4682
110. Dong L, Hu S, Gao J. 2020. Discovering drugs to treat coronavirus disease 2019 (COVID-19). *Drug Discov. Ther.* 14:58–60
111. Takahashi H, Iwasaki Y, Watanabe T, Ichinose N, Okada Y, et al. 2020. Case studies of SARS-CoV-2 treated with favipiravir among patients in critical or severe condition. *Int. J. Infect. Dis.* 100:283–85
112. Dietzschold B, Wunner WH, Wiktor TJ, Lopes AD, Lafon M, et al. 1983. Characterization of an antigenic determinant of the glycoprotein that correlates with pathogenicity of rabies virus. *PNAS* 80:70–74
113. Salvato M, Borrow P, Shimomaye E, Oldstone MB. 1991. Molecular basis of viral persistence: A single amino acid change in the glycoprotein of lymphocytic choriomeningitis virus is associated with suppression of the antiviral cytotoxic T-lymphocyte response and establishment of persistence. *J. Virol.* 65:1863–69
114. Sitbon M, d'Auriol L, Ellerbrok H, Andre C, Nishio J, et al. 1991. Substitution of leucine for isoleucine in a sequence highly conserved among retroviral envelope surface glycoproteins attenuates the lytic effect of the Friend murine leukemia virus. *PNAS* 88:5932–36
115. Bae YS, Yoon JW. 1993. Determination of diabetogenicity attributable to a single amino acid, Ala776, on the polyprotein of encephalomyocarditis virus. *Diabetes* 42:435–43

116. Holzmann H, Stiasny K, Ecker M, Kunz C, Heinz FX. 1997. Characterization of monoclonal antibody-escape mutants of tick-borne encephalitis virus with reduced neuroinvasiveness in mice. *J. Gen. Virol.* 78(1):31–37
117. Núñez JI, Baranowski E, Molina N, Ruiz-Jarabo CM, Sánchez C, et al. 2001. A single amino acid substitution in nonstructural protein 3A can mediate adaptation of foot-and-mouth disease virus to the guinea pig. *J. Virol.* 75:3977–83
118. Nakano Y, Yamamoto K, Ueda MT, Soper A, Konno Y, et al. 2020. A role for gorilla APOBEC3G in shaping lentivirus evolution including transmission to humans. *PLOS Pathog.* 16:e1008812