Preprints.org

Quasispecies Theory and Emerging Viruses: Challenges and Applications

Josep Sardanyes [*] , Celia Perales , Esteban Domingo , Santiago Elena

Review

Quasispecies Theory and Emerging Viruses: Challenges and Applications

Josep Sardanyés [1,2,*], Celia Perales [3,4], Esteban Domingo [5] and Santiago F. Elena [6,7]

[1] Centre de Recerca Matemàtica, Edifici C, Campus de Bellaterra, Cerdanyola del Vallès, 08193 Barcelona, Spain

[2] Dynamical Systems and Computational Virology, CSIC Associated Unit I2SysBio-CRM

[3] Department of Molecular and Cell Biology, Centro Nacional de Biotecnología, CSIC, Cantoblanco, 28049 Madrid, Spain

[4] Department of Clinical Microbiology, Instituto de Investigación Sanitaria, Fundación Jiménez Díaz University Hospital-Universidad Autónoma de Madrid, 28040 Madrid, Spain

[5] Microbes in Health and Welfare Program, Centro de Biología Molecular "Severo Ochoa", CSIC-Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

[6] Instituto de Biología Integrativa de Sistemas (I2SysBio), CSIC-Universitat de València, Paterna, 46980 València, Spain

[7] The Santa Fe Institute, Santa Fe, New Mexico 87501, USA

* Correspondence: jsardanyes@crm.cat

Abstract: Quasispecies theory revolutionized our understanding of viral evolution by describing viruses as dynamic populations of genetically diverse variants constantly adapting. This article explores the theory's role in virus-host interactions, including immune evasion, drug resistance, and viral emergence. We review the original model, recent advances, and key virus dynamics needing incorporation into quasispecies theory. We introduce the *ultracube* concept as a more realistic multidimensional sequence space to investigate virus evolutionary dynamics.

The quasispecies theory, conceived by Manfred Eigen and Peter Schuster more than fifty years ago [1–4] was developed to investigate the dynamics of biological information in replicators subjected to exceptionally high mutation rates. This theory is a cornerstone in understanding prebiotic evolution and a new framework to investigate the genetic diversity of viruses and their dynamics [5–10]. Moreover, it also extends to the dynamics of cancer cells [11–13] and has been linked to the conformational diversity of prions [14]. At its core, the theory posits that viral populations exist not as static entities with a single genome, but rather as dynamic distributions of closely related mutant genomes known as mutant swarms or *quasispecies*. The term diversity refers to genetic and phenotypic differences among clades or isolates of viruses from the same taxonomic group, or of the viral world (virosphere) in general. In contrast, quasispecies refers to a particular population structure that describes the genome diversity within a virus isolate or a laboratory population, its dynamical properties (including mutational coupling between genetic variants), and its consequences for virus evolution and pathogenesis. Viral populations are characterized by high mutation rates due to the limited template-copying fidelity of RNA-dependent RNA polymerases (RdRp) and RNA-dependent DNA polymerases (RdDp), leading to the continuous generation of genetic variants. The collective behavior of these variants, along with multiple selection pressures—such as heterogeneity in susceptible target cells, host immune systems, and potential antiviral therapies—and bottleneck events at various scales, together shape the evolutionary trajectories of the virus. Thus, understanding the quasispecies dynamics is essential for elucidating viral pathogenesis, transmission dynamics, and the emergence of drug resistance. The molecular quasispecies theory has driven the development of a comprehensive theory of virus evolution, supported and enriched by numerous experimental and clinical studies. Figure 1 shows a brief outline of some key contributions

to viral quasispecies. The present article is intended for both theoreticians and experimentalists, and aims to provide a knowledge bridge between these audiences and both clinical and medical researchers. We review in an interdisciplinary manner the major historical developments of quasispecies theory, and new concepts that have arisen as the result of mutual influences between advances in experimental observations and theoretical aspects. We delineate some challenges in biomedicine related to the impact of quasispecies dynamics on disease control and the emergence of new viral pathogens. These points may also interest the medical community involved in investigations of viral disease mechanisms and therapeutic approaches.
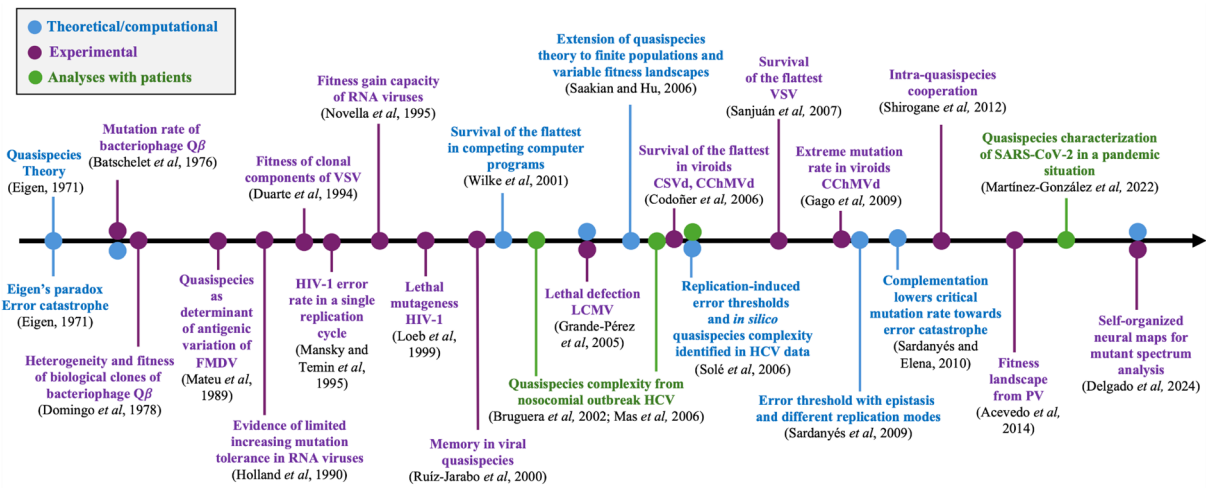

Figure 1. Milestones in quasispecies theory in Virology. Time arrow showing some key achievements in the development of quasispecies theory for viruses since its birth in 1971. The colors indicate whether the results have been achieved from theoretical/computational research (blue), from experimental data (violet), and/or from data from infected patients (green). CChMVd: Chrysanthemum chlorotic mottle viroid; CSVd: Chrysanthemum stunt viroid; FMDV: foot-and-mouth disease virus; HCV: hepatitis C virus; HIV-1: human immunodeficiency virus type-1; LCMV: lymphocytic choriomeningitis virus; VSV: vesicular stomatitis virus; PV: poliovirus.

Research on bacteriophage Qβ by Charles Weissmann and colleagues in the 1970s [15] coincided with the development of quasispecies theory by Eigen and Schuster; the two investigations were carried out independently. During the 1980s and 1990s, scientists compared genomic sequences of clones from natural viral isolates or experimental viral populations. The introduction in the 2000s of ultra-deep sequencing significantly advanced the understanding of the genetic complexity of the mutant swarm. The quasispecies concept has profound implications for clinical and public health interventions. By viewing viral populations as dynamic ensembles of genetic variants rather than static entities, the quasispecies theory underscores the challenges to eradicating viral infections through conventional interventions. Antiviral therapies that target a single viral genotype may exert a selective pressure, favoring the emergence of drug-resistant mutants within the cloud of mutants forming the quasispecies. This evolutionary resilience necessitates the development of novel therapeutic strategies that take into account the complex evolutionary dynamics of viral populations. Additionally, insights from the quasispecies theory have led to advancements in experimental and computational techniques for characterizing viral diversity. These improvements allow for more precise monitoring of viral evolution and help design programs to reduce the spread of viral pathogens.

Two related key concepts in quasispecies theory are the so-called sequence spaces and fitness landscapes. The sequence space is a multidimensional discrete space, also called a hypercube, where each node corresponds to a given genotype, which is connected to the neighboring genotypes by single-point mutations [16] (Figure 2a). Viral fitness is the capacity of a virus to produce infectious progeny. It is often expressed as a relative value, taking as reference a population or clone of the same virus. It is an environment-dependent value, so it need not be identical when measured with different

cell lines in culture, organoids, tissues, organs, or an entire live host. Its value depends on multiple host and viral factors, including the mutant spectrum composition. The term "epidemiological fitness" was coined to refer to differences among viral strains (or clades, lineages, variants) to become dominant in the epidemiological setting where parameters other than replicative capacity (*i.e.*, particle stability, transmissibility) also play a role. Because of these multiple influences, viral fitness varies with time. Also, the need to understand the structure of fitness landscapes [17] applies to viral populations (different aspects of fitness have been reviewed in Refs [18,19].
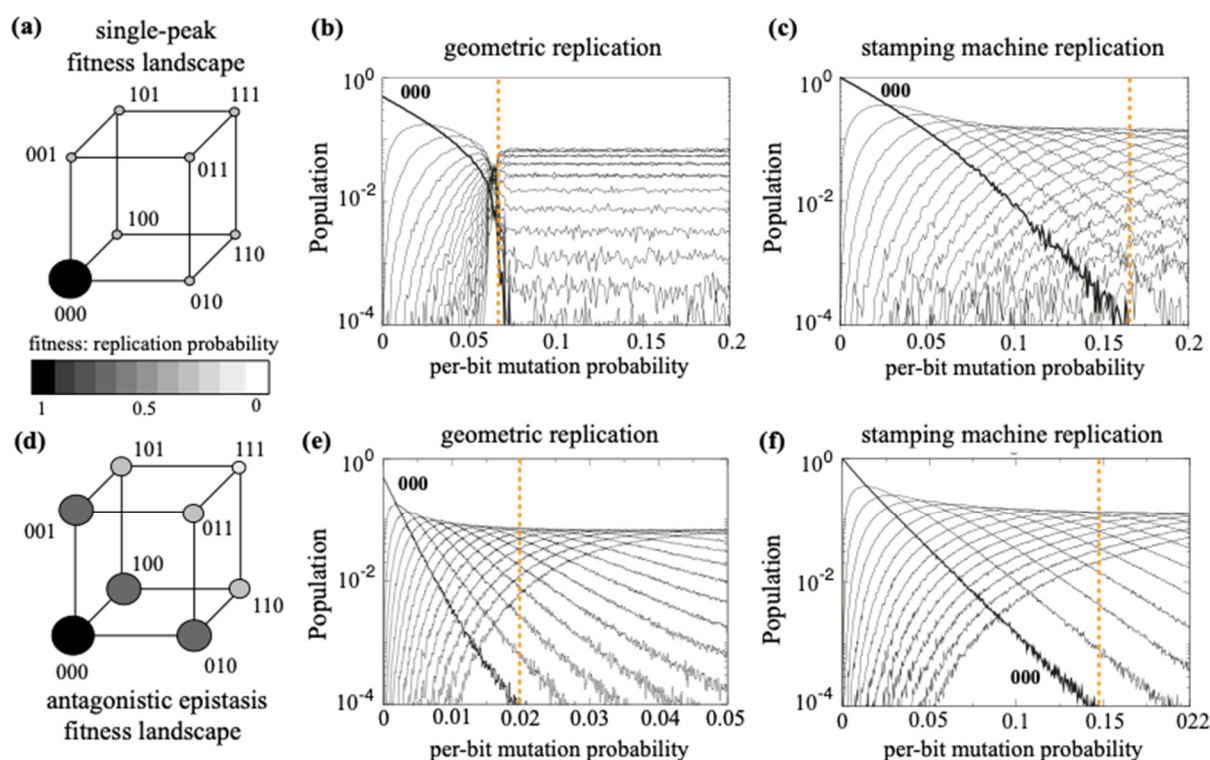

Figure 2. Simulating the evolutionary dynamics of digital quasispecies *in silico*. (a) Single-peak fitness landscape (illustrated with a 3-bits hypercube).The size of the balls is proportional to the genotypes' fitness.   The quasispecies population plots show the error threshold for the single-peak fitness landscape for geometric (b) and stamping machine (c) replication. (d) Fitness landscape with antagonistic epistasis also for 3-bit genomes where the effects of mutations are less severe in combination than individually. Here we also display the error threshold for e for geometric (e) and stamping machine (f) replication modes. All the diagrams show the stationary populations of the master sequence (000, thick lines) and the pool of mutants (with 1, 2 or 3 mutations, thin lines) averaged over 200 replicas at increasing the per-bit mutation probability. See Sardanyés et al. (2009) for further details.

Generically, the order of a hypercube is n^L, n being the number of letters in the alphabet (4 for nucleotides) and L the length of the genome (which determines the dimension of the hypercube). That is, the order of the sequence space for an RNA virus with a short genome, such as bacteriophage MS2, with 3569 nucleotides, results in a hypercube with 4^{3569} connected nodes. This is a huge sequence space that can be explored by the virus through the processes of point mutations or recombination events that do not alter the genome length (Figure 3a). Each of the nodes of this hypercube thus corresponds to a different genotype with a different fitness. It is known that point mutations impact the fitness of the viral genotypes, typically being deleterious or lethal [20–25]. The quasispecies is expected to be located in a given region or regions of this multidimensional space after selection takes place. As we discuss below, these hypercubes may not accurately describe the complexity of quasispecies when considering other genetic processes, such as deletions or insertions, which result in viral genomes of varying lengths. Hence, the quasispecies should be expected to live in a much more complex sequence space that we define as the *ultracube* and will explain below.

The distribution of fitnesses along the hypercube nodes defines the so-called fitness landscape. A fitness landscape is a conceptual model used in evolutionary biology that metaphorically represents the multidimensional space where each point corresponds to a specific genotype and is associated with a quantitative value that represents its fitness. For example, in the most visually appealing two-dimensional representation, the coordinates in the plane represent genotypes, and the elevation over the surface indicates the fitness of that genotype. In simpler terms, a fitness landscape summarizes how well an organism with a particular genotype is adapted to its environment and, *e.g.*, performs in replication. Fitness landscapes can vary in complexity, ranging from simple, smooth landscapes with a single peak representing the optimal genotype to rugged landscapes with multiple peaks and valleys, indicating the presence of different adaptive solutions or evolutionary pathways [16,26,27]. The shape of the fitness landscape is influenced by several factors, including the genetic architecture of traits, the nature of environmental conditions, and the interactions among different genotypes. Having a precise fitness landscape of viral quasispecies is an extremely challenging problem. For real viruses, fitness landscapes are increasingly viewed as very rugged and dynamic [19,28–31].

Theoretical quasispecies: what has been explored and existing gaps

The population dynamics of quasispecies have been extensively studied since the original contributions by Eigen and Schuster. Several mathematical and computational models have been used to investigate various aspects of viral quasispecies. Here, we discuss the main contributions of quasispecies theory to the field of RNA viruses, highlighting several processes that have not yet been investigated in detail from a theoretical or computational point of view and may be worth exploring. The original Eigen-Schuster quasispecies model is given by the set of differential equations

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} x_j f_j Q_{ji} - \Omega(x) x_i.$$

This mathematical model describes the time change of the fraction of the population of the ith mutant sequence x_i $(i = 1, \ldots, n) \in x$, and n being very large. Here f_j is the replication rate of the jth mutant, Q_{ji} is the probability of having a mutation $j \rightarrow i$, and $\Omega(x) = \sum_{j=1}^{n} f_j x_j$ denotes the average fitness of the population. One key aspect of quasispecies models is that the sequences inhabit a fitness landscape; in this case, each sequence has a given fitness value, which determines its rate of replication. An infecting wild-type (wt) virion can produce an enormous progeny inside the host, giving rise to a large amount of different sequences embedded in a more or less complex fitness landscape. Despite such complexity, the quasispecies model can be explored in a simpler setting, allowing to mathematically characterize phenomena of interest that may also occur in more complex scenarios.

One such useful simplification is to consider a quasispecies formed by only two populations of genomes: the wt and the pool of mutants. This simple model easily illustrates a fundamental consequence of quasispecies theory: the error catastrophe or error threshold. This two-populations model assumes that the quasispecies is embedded in a single-peak fitness landscape [3,11,32,33] and the wt sequence, x_0 with high fitness f_0, produces deleterious mutants that are grouped into an average mutant sequence, x_1, all with equal fitness $f_1 < f_0$. This model also assumes that backward mutations are negligible. This is a first good assumption due to the enormous size of the sequence space. This model can be written taking as variables these two sequence types, with:

$$\frac{dx_0}{dt} = f_0 x_0 (1 - \mu) - \Omega(x_0, x_1) x_0 \, ,$$

$$\frac{dx_1}{dt} = \mu f_0 x_0 + f_1 x_1 - \Omega(x_0, x_1) x_1 \, .$$

Here μ is the mutation rate, and $\Omega(x_0, x_1)$ is an outflow term given by the average fitness and which keeps a constant population (another unrealistic assumption for real viral quasispecies). This oversimplified model allows the calculation of the error threshold occurring when mutation overcomes the critical value $\mu_c = 1 - f_1/f_0$. When $\mu < \mu_c$ the quasispecies is composed by both the wt and mutant sequences, while for $\mu > \mu_c$ the entire quasispecies is only composed of mutants. In

other words, this critical mutation marks a transition between a phase wherein the genetic information carried by the wt sequence is preserved and a phase in which this information is no longer maintained [1,32–36]. Interestingly, more complex versions of the model that incorporate epistatic fitness landscapes and different modes of genomic replication (*e.g.*, geometric versus stamping machine amplification) still preserve the existence of an error threshold, although its actual value strongly depends on these assumptions and can be even bigger under realistic combinations of parameters (Figure 2), as we discuss below. For instance, panels (b) and (c) in Figure 2 show that the critical mutation rate involving a full dominance of mutants is much higher under the stamping machine mode, meaning that the master sequence is preserved in a wider range of mutation rates, as compared with the geometric mode. Moreover, the change in the fitness landscape (how mutations impact a fitness trait, *i.e.*, genomes' replication speed) has also a big impact on the distribution of the master sequence and the cloud of mutants. A fitness landscape qualitatively mimicking antagonistic epistasis [Figure 2d] under the geometric replication mode displays an extremely small critical mutation rate. This lower critical value probably arises because mutants close to the master sequence still have good fitness (lower than the master sequence) and thus competition promotes the extinction of the master sequence at $\mu_c = 0.02$ [Figure 2e]. This critical value changes to $\mu_c = 0.15$ for the stamping machine replication [Figure 2f].

Moreover, oversimplifications allowing for mathematical modeling are sometimes meaningful. For instance, research on quasispecies in HCV-infected patients revealed an inverse correlation between viral load and quasispecies complexity [7]. Such a feature was reproduced with the two-populations model described above using *in silico* quasispecies evolution [33].

The original Eigen-Schuster quasispecies model was built upon several assumptions that might not hold for real viruses. It assumes continuous, well-mixed populations of replicators, constant population, geometric replication, and determinism. During the last decades, considerable efforts have been made to extend the initial quasispecies theory to more realistic scenarios for RNA viruses. Examples comprise the study of other key features of viral quasispecies that are summarized next and which have been investigated separately or in combination. These include finite populations [12,37,38], stochastic effects [33,38,39], spatially-extended quasispecies [38,40–42], viral complementation [35], and recombination [43,44]. The investigation of asymmetric modes of replication, identified in RNA viruses either indirectly from mutant distributions [45–47] or from direct RNA quantification of positive and negative strands during the progress of infection [48,49], have been also studied within the quasispecies framework [42,50].

Another relevant theoretical result of quasispecies theory is the so-called survival-of-the-flattest effect. This effect is mainly produced because fast replicating genomes that produce low-fitness offspring can be outcompeted by slow replicating genomes with moderate fitness, provided the latter inhabit a region of sequence space characterized by high neutrality and connectivity [51,52]. This theoretical prediction was first demonstrated in artificial life experiments [53] and later described in experiments with competing viroids [54] and withVSV [55] under mutagenic conditions. This effect was later explored mathematically, and the out-competition of the fit quasispecies by the flat one at an increasing mutation rate was shown to be given by an abrupt transition [38]. Flat-like quasispecies may underlie failures in resolving chronic infections by antiviral agents, despite the absence of specific inhibitor-resistance substitutions [56]. Concerning the fitness landscapes, the Swetina-Schuster single-peak one is, of course, an oversimplification of how mutations impact fitness in viral quasispecies. During the last decades, more complex fitness landscapes have been studied for quasispecies [27]. These include different fitness functions [57], dynamic fitness landscapes [58,59] and non-linear interactions among mutations, *i.e.*, epistasis [9,42,50].

As we have discussed, the original quasispecies theory has been significantly refined over the past decades to better align with real virology. Relevant experimental and clinical results involving RNA viruses have been interpreted through the lens of these theoretical models. Despite this fact, there are still missing ingredients in quasispecies theory that could play a crucial role in the evolutionary dynamics of viruses. For example, some processes, such as RNA viral synthesis or viral protein production and maturation, involve time lags. In this sense, several studies have revealed

that the SARS-CoV-2 replication complex has an elongation rate of 150 to 200 nucleotides per second, being more than twice as fast as the poliovirus polymerase complex [60]. Nucleotide incorporation time, determined by differences at and around viral RNA polymerase catalytic sites, may influence the fidelity of template copying [61]. In this sense, it is known that time lags can profoundly affect the dynamics of nonlinear systems, causing self-sustained oscillations and chaos [62]. Also, during their intracellular phase, viruses synthesize distinct molecules, constructing viral factories for genome replication and morphogenesis. The central element of these factories is the replication organelle (RO), where viral replication complexes produce multiple copies of the viral genome. Viral factories often consist of remodeled cell membranes with functional compartments for replication, assembly, and egress, and they frequently recruit other cellular elements like mitochondria and the cytoskeleton, which interact with the RO. Unlike DNA viruses, almost all RNA viruses form factories exclusively in the cytoplasm [63,64]. For instance, HCV induces remodeling of reticulum endoplasmatic membranes forming double-membraned vesicles, and later on, it induces the formation of multi-membrane vesicles that are composed of several concentric membrane bilayers. These membranous rearrangements are produced by the action of viral nonstructural proteins. Hence, the process of virus genome amplification can be highly compartmentalized inside the cell and possibly also induce time lags in viral replication and virion assembly.

For RNA viruses, especially those infecting plants, replication processes may exhibit periodic fluctuations at the within-tissue or within-host levels across different time scales, primarily due to temperature changes [65]. Additionally, the mammalian brain possesses an endogenous central circadian clock that regulates both central and peripheral cellular activities. At the molecular level, this day-night cycle triggers the expression of upstream and downstream transcription factors that affect the immune system and modulate the severity of viral infections over time. The role of circadian systems in regulating viral infections and the host response to viruses is thus of great clinical importance [66], and thus fluctuating parameters may be included in theoretical quasispecies investigations.

Another unexplored process in quasispecies theory is how the evolutionary dynamics of quasispecies at the within-cell/within-host levels scale up to the population-epidemiological levels. This question, which we discuss below, can be addressed by developing multi-scale models including the rapid quasispecies evolution (fast dynamics) and the impact of the continuous synthesis of heterogeneous mutants at a population scale (slow dynamics). The mathematical results on slow-fast systems may be thus relevant to tackle this problem. In this sense, the connection of scales in virology is extremely important to understand virus pathogenesis and potential zoonoses and epidemic spread among different hosts [36]. This subject can be further explored combining dynamical systems theory with complex networks theory.

The limits of predictability of quasispecies are also an extremely challenging and open problem in virology. That is, how the dynamical population structure of the quasispecies may impact disease outbreaks and the emergence of new variants with epidemic and pandemic potential (see Section on **Impact of quasispecies populations in emerging viral diseases** below).

Quasispecies and RNA viruses: genomic heterogeneity, adaptation and clinical implications of genetic information thresholds

Initial experimental evidence of viral quasispecies dynamics involved studies mainly with bacteriophage Qβ, VSV, FMDV, or LCMV (reviewed in Ref. [10]). These experiments have robustly confirmed high mutation rates during genome replication, the heterogeneity of viral populations, fitness variations among biological clones, rapid changes in sequence space occupation, and limited tolerance to increased error rates, to cite a few. The early studies with bacteriophage Qβ by Weissmann et al. were carried out simultaneously with the development of quasispecies theory by Eigen and Schuster in the 1970s. In the 1980s and 1990s, the primary approach was to compare genomic sequences of molecular or biological clones from natural viral isolates or viral populations subjected to various experimental evolution designs. The advent of ultra-deep sequencing techniques greatly broadened the capacity to probe the complexity and dynamics of viral quasispecies. The extended capacity to probe into the complexity and dynamics of viral quasispecies afforded by ultra-

deep sequencing is exemplified by results on HCV [7,67] and, more recently, in SARS-CoV-2 displaying multitudes of low-frequency mutations in isolates of the virus, some of them endowed with functional relevance [68–70].

During recent years, the development of new high-throughput sequencing techniques (*e.g.*, Cir-Seq) along with novel bioinformatic algorithms have shown the relevance of an abundant component of viral quasispecies that was previously largely ignored: the defective viral genomes (DVGs) [71–74]. Broadly speaking, DVGs are nonstandard genomes generated during error-prone replication that contain deletions, insertions, duplications, inversions, and potential hypermutated viral genomes contributed by cellular editing activities. These major-effect mutations render genomes unable to self-replicate, and coinfection with a wt (also dubbed helper) virus is needed for the defective genomes to persist in the population. Several recent studies have described the temporal dynamics of DVGs during the course of infection for several viruses, including poliovirus (PV) and dengue virus (DENV) [75], influenza A virus (IAV) [76], SARS-CoV-2 [77] and other betacoronaviruses [74]. Consistently, these studies show that some DVGs are pervasively maintained in the viral population in cell cultures but also in patients, suggesting a possible selective role. For instance, deletions affecting the receptor binding domain and the S1/S2 cleavage site in SARS-CoV-2 Omicron variant may have increased host cell ACE2 receptor recognition, thus enhancing the infection and allowing this variant to become dominant [78]. Indeed, it has been suggested that DVGs might confer some advantage to viruses, such as serving as reservoirs of genetic variability, decoys for immune responses, regulators of translational shut-down, or mediators of persistent infections [72].

A particularly notable type of DVG due to its structural and GC content requirements is the copy-back and snap-back (cb/sb) variants, which consist of small RNA molecules with a hairpin-loop structure created by template-switching from positive to negative templates during replication.

As we have mentioned, an increase in population mutational load should concomitantly result in a decrease in the average population fitness as most mutations with phenotypic effects are deleterious or lethal [8,20,22,25,73]. Given this, and together with the aforementioned prediction of the existence of a critical mutation rate beyond which the population enters the error catastrophe regime [1,4], innovative antiviral strategies have been proposed aiming to push down replication fidelity and forcing the viral population to cross the error threshold. Nucleotide analogues can act as antiviral agents, and one of their mechanisms of activity is the elevation of the viral polymerase error rate beyond the maximum value compatible with maintenance of viral infectivity [80,81]. This strategy adds to others intended to minimize selection of escape mutants, for example, sequential antiviral administrations, use of antiviral agents that target cellular functions which are needed for completion of the virus infectious cycle, combination of immunotherapy with chemotherapy, use of antiviral agents that enhance the innate immune response, etc. These different approaches have been investigated experimentally and with theoretical models that computed the risk of antiviral escape and means to counteract treatment failures [82–85] (see also Ref. [86] and references quoted therein).

Quasispecies complexity: from the *hypercube* to the *ultracube*

In the study of viral evolution and coevolution, hypercubes provide a useful conceptual framework for comprehending the sequence space of viral genomes [4,27,87]. This sequence space is a mathematical multidimensional construct where each dimension corresponds to a nucleotide position within the viral genome (Figures 2 and 3a). A hypercube, in particular, is a discrete geometric representation of this space, with vertices representing every possible sequence configuration in which a viral genome of length L can exist. Each vertex is connected to adjacent vertices by edges that signify single nucleotide changes, illustrating the potential mutational pathways a virus can traverse while retaining the same length. This model effectively visualizes the extensive diversity and evolutionary potential of viral populations. The structure of the hypercube allows researchers to analyze the distribution and dynamics of viral quasispecies within this space, providing valuable insights into how viruses adapt to environmental pressures and develop resistance to antiviral treatments. The foundational works of Eigen and Schuster [1,2] were instrumental in introducing the concepts of sequence space and hypercubes within the framework of quasispecies theory, offering a

mathematical basis for understanding the evolutionary landscape of viruses [3]. However, viral quasispecies may be embedded within more complex and entangled sequence spaces, including the full spectrum of mutant sequences together with, *e.g.*, subgenomic sequences and DVGs. These length variants may form other subpopulations of clouds of mutants spanning lower- or higher-dimensional hypercubes arising from the full-genome sequence space, thus being connected through deletions or insertions. Hence, a more realistic geometric space for viruses may include connected hypercubes of different dimensions: we call these sequence spaces, in which a given node can represent a hypercube itself, as *ultracubes* (Figure 3b). This view enlarges the size of the sequence space beyond the single-point mutation hypercube dimension. As for the hypercube, a quasispecies may not span all this sequence space but be located on some specific regions of this ultracube in a mutation/recombination/deletion-selection balance.
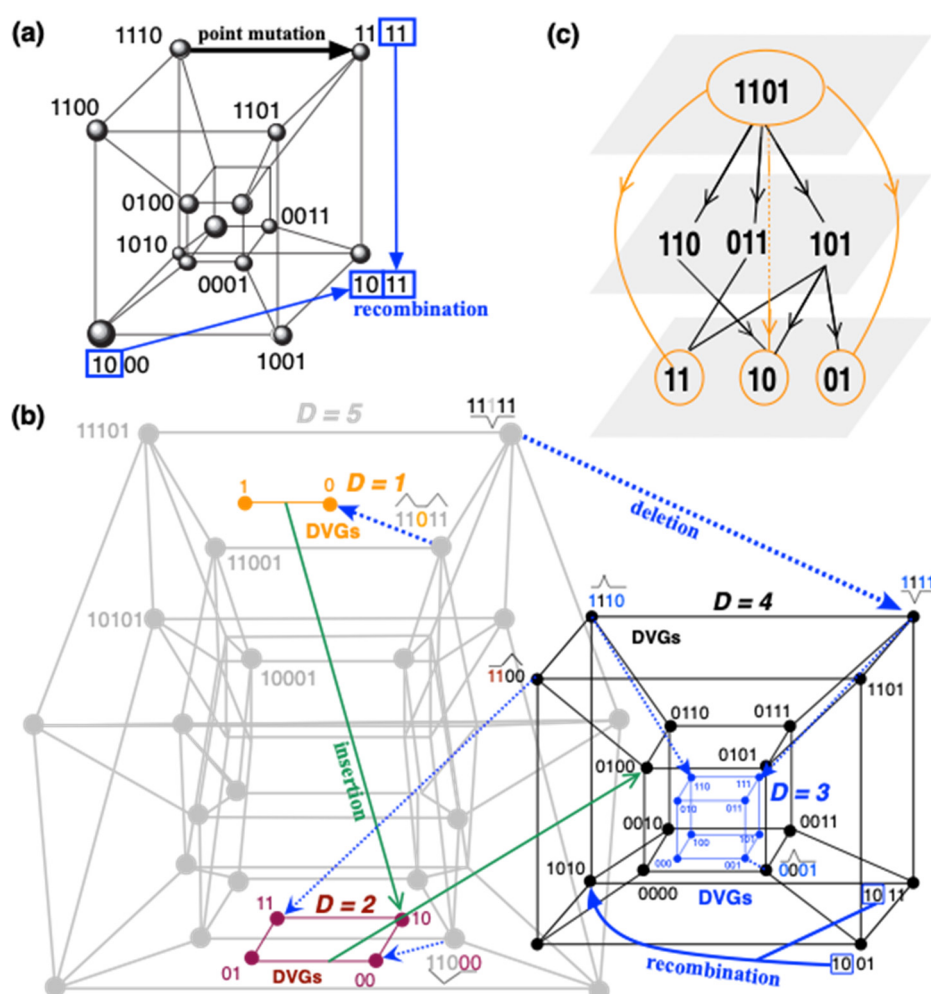

Figure 3. Virus genotypic complexity: from the *hypercube* to the *ultracube*. (a) Classical view of quasispecies evolution in a hypercube for 4-bit sequences. Each node of the network connects two genotypes via a single-point mutation. Sequences evolve to first neighbors by single-bit (nucleotide) substitutions during replication. Homologous recombination may allow genotypes to jump to further neighbors (blue arrow). (b) Example of a sequence space for binary genomes of length five considering deletions (blue dashed arrows) and insertions (green solid arrows) during replication. These processes produce mutants and connect hypercubes of different dimensions, giving rise to a more complex sequence space that we label *ultracube* and which can be conceived as a multiplex network. For clarity, we do not display all the nodes but exemplify some processes of deletion and insertion, which give rise to a set of connected hypercubes of dimensions 5 (gray), 4 (black), 3 (blue), 2 (red),

and 1 (orange). (c) Schematic diagram of connected hypercubes of different dimensions illustrated schematically as multilayered networks.

Multilayer models for multi-scale virus dynamics: integrating quasispecies into virus epidemiology and ecology

A new area of complex networks theory that has been quickly developing over the last decade is the study of multilayer networks. Each layer contains nodes connected by intralayer edges that describe rules of interactions between the nodes for this particular layer. In addition, dependencies across layers are represented by interlayer edges. It is straightforward to place and visualize viruses into a multilayer; for illustrative purposes, let us discuss here a simple 3-layer case. The bottom layer would represent the quasispecies generated within a single infected individual. At this level, the ultracube (Figure 3b) would be an appropriate representation. The equations describing the dynamics at this bottom level could be those described in the previous sections incorporating mutants of different nature. A middle layer would represent local host populations. At this level, each node corresponds to a particular host, and the edges represent the contact network among the hosts that determines transmission dynamics. The equations governing the virus' dynamics at this intermediate level could be the well-known SIR epidemiological model, and the network can show any topology, such as scale-free [88]. The interlayer edges connecting the bottom and the middle layers represent the probability that a given viral genotype is actually present in each infected host, which indeed depends on the replicative advantage at the bottom level and the mutation rate. Coinfections are allowed (two interlayer edges pointing towards the same individual). The upper layer represents the epidemiological level. In this layer, nodes represent, *e.g.*, communities, whereas edges represent the connectivity between these communities (*e.g.*, airport traffic connections, vector flights, pollen and seed dispersal, etc.). The dynamics at this level can be modeled using phylogeography tools. The interlayer edges connecting the middle and the upper layers would represent the probability that an infected individual will move from one community to another.

This multilayer representation allows us to study not only the dynamics at each layer but also the entire multilayer system and infer properties such as multilayer modularity, robustness to perturbation or phase transitions. The intra-host evolutionary dynamics (bottom layer) has been modeled in connection with epidemic spread (upper layer) [89]. Viral genotypes closely connected in the quasispecies bottom layer may be found in individuals in the middle layer that also form a transmission cluster (or module). For instance, ecological multilayer networks for plant-aphid and plant-aphid-parasitoids [90] show nontrivial stability properties that result in quantitative predictions about the persistence or extinction probabilities that would not be shown up by other modeling approaches [91]. Information between nodes in a particular layer can be transmitted via two different paths: those involving only intralayer edges and those involving both intra- and interlayer edges. This means that catastrophic failures in a particular layer can easily be transmitted to the rest of layers [92,93]. An example of catastrophic failures across levels in biomedicine is the understanding of pathological conditions via the network cell organization into genes, proteins and metabolites. Cell malfunctions are rarely the result of one of these three levels. Rather, they result from multiple interactions at the three levels. This view reinforces the prospects for personalized medicine whose impact in viral disease control is still to be investigated. This, and other examples in different domains of science, of the consequences of inter-layer interactions are described in Ref. [93]. The epidemic spread of infectious diseases in multilayer networks has received some attention [94–97]. Examples are: (*i*) In particular, a topic that has been studied is the interaction between two genetic variants of the same virus, showing that coexistence or displacement of the less fit variant depends on the connectivity between layers [96,98–100]. (*ii*) In multilayer networks, the disease can spread through multiple channels or layers simultaneously, which affects the overall dynamics. For instance, a disease might spread rapidly in one layer (*e.g.*, in the example of vector-borne viral diseases, within the vertebrate host population) and more slowly in another (*e.g.*, in the same example, within the insect population) [91,99]. The coupling between these layers can either enhance or inhibit the spread, depending on factors like the strength of connections between layers or the level of interaction between them. (*iii*) In classic epidemiological SIR or SEIR models, there are often R_0 thresholds that

determine whether an epidemic will occur. In multilayer networks, these thresholds became more complex. The disease might die out in one layer (*e.g.*, in our example of the vector-borne virus, from the vertebrate population) but continue spreading in another (*e.g.*, in the mosquito reservoir), or it might require a critical level of interaction across layers before an epidemic can fully develop [88]. (*iv*) Multilayer networks naturally introduce heterogeneity into the system. Different individuals or nodes may have different degrees of connectivity in each layer, leading to varied exposure risks. For example, a person who is highly connected in one layer (*e.g.*, traveling overseas a lot) but less connected in another (*e.g.*, in its home local community) may play a unique role in disease transmission [73,101]. Finally, (*v*) fitness values are environment- and population context-dependent. Positive correlation of fitness values across the intra-host and the inter-host (transmission) scales can facilitate a successful pathogen emergence [102]. A stochastic model further determined that a wide transmission bottleneck can help the emergence, unless pathogens exhibit cross-scale selective conflicts [103].

When multiple variants are spreading in a host scale-free network and competing for hosts, they obviously influence each other's dynamics. In such a situation, for example, the derived epidemic threshold for each variant is substantially different than predicted using monolayer simple coinfection models. Indeed, two new thresholds arise from these models [99]: (*i1*) the survival threshold determines a continuous phase transition from extinction to existence during competition between both variants. (*ii2*) The absolute-dominance threshold denotes the critical point where one of the variants fully outcompetes the other. Between these two thresholds, coexistence emerges as a property of the interconnected structure of the multilayer. The incorporation of quasispecies models within multilayer networks may allow considering the impact of the microevolutionary dynamics of viral quasispecies at a multi-scale level. In this sense, continuously generated variants may impact at a populational level (epidemiological scale), and phenomena such as the error threshold could determine transitions in upper layers. A discussion on the multi-scale nature of viruses is discussed in Ref. [36].

Impact of quasispecies populations in emerging viral diseases

The emergence of viral disease is an unpredictable event since it is the result of several interconnected factors: the adaptive potential of viruses circulating in different host species, environmental modifications, or political and sociological circumstances, among others. These factors can affect virus traffic and facilitate encounters with potential new hosts [104–106]. Viral emergence can be regarded as a facet of complex biological behavior since its occurrence cannot be anticipated by the sum contribution of the multiple underlying factors [36,107,108]. Since quasispecies enhances viral adaptability and microbial adaptability is one of 13 factors involved in transmissible disease emergence [106], it is sometimes assumed that there is a direct connection between quasispecies and viral disease emergence. A direct connection is difficult to prove experimentally. However, several studies have shown that the presence of cell tropism and host range mutants in mutant spectra (or the presence of an ample mutant repertoire) mediated the adaptation of a virus to a different host species. Examples include parvoviruses, coxsackievirus B and foot-and-mouth disease virus (reviewed in Ref. [109]). In this scenario, the production of mutant viruses at high frequency in immunocompromised individuals [described for several viruses [109] including SARS-CoV-2 [110,111] may represent an additional supply of potential host range modifications. All the genotypic complexity portrayed by the ultracube connections (Figure 3) adds to the lottery of which variants (and when) they may come into contact with a susceptible individual of a different host species, to initiate an infection and to produce sufficient transmissible progeny to achieve the status of emergent disease.

John Holland and colleagues were the first to point out a number of medical implications of a highly dynamic RNA world coexisting with a relatively more static DNA-based cellular biosphere [112]. They underlined high mutation rates as a source of atypical viral forms capable of invading new tissues and organs, to establishing persistent or inapparent infections, or beingt o be selected in response to medical interventions. These facets of viral dynamics have been amply documented in the 1940s that have followed their prescient publication. Joshua Lederberg warned of the human

vulnerability in the face of the adaptive potential of RNA viruses: "Abundant sources of genetic variation exist for viruses to learn new tricks, not necessarily confined to what happens routinely or even frequently" [113]. In May 2003, the World Health Organization released a report titled "SARS: Status of the outbreak and lessons for the immediate future" [114]. Work on population heterogeneity and quasispecies dynamics of the emergent SARS-CoV-2 has confirmed coronaviruses as a genetically and functionally highly variable virus group, and as a threat to produce future emerging diseases [31].

Future challenges in experimental and clinical quasispecies

As often happens in science, new techniques pose new questions. The availability of deep sequencing methodology to probe into the composition of virus populations has opened new questions, in addition to solving old ones such as the definitive confirmation that mutant spectra are the biological reality of viral populations, so far without documented exceptions. Among the new challenges are: (*i*) capturing the depth of minority genomes that until now have escaped detection; (*ii*) how they feed the dominant ones, be them viable or defective; (*iii*) how the latter modulate behavior of the ensemble; (*iv*) which is the time frame in which minority subpopulations can be replaced by others, relative to the time required for selective forces (metabolic modifications, signal effectors, antibodies, drugs, etc.) to reach the extracellular and intracellular environments or even individual viral factories; (*v*) what is the extent of heterogeneity within individual cells; and (*vi*) capacity of mutant spectrum composition to inform of long-term evolutionary events. These are but some of many highly relevant questions. What is at stake is to better understand viruses, the diseases they produce, and the means to combat them. In short, a challenge is to shift from a consensus sequence-centered understanding of evolution into a mutant spectrum-centered understanding of evolution, and harmonizing the conclusions drawn from both.

Therapies based on the error catastrophe concept have been brought forward [81,115]. Maintenance of inheritable genetic information conditioned to a limitation in error introduction has also been documented with catalytic ribozymes [116]. Lethal mutagenesis approaches exploit the intrinsic balance between mutation rates and viral viability, offering a promising avenue for combating viral infections. As a difference from the error threshold, which causes a shift in the sequence space, *i.e.*, the master sequence is replaced by the mutant swarm, lethal mutagenesis involves virus extinctions [32]. Virus extinction by lethal mutagenesis has been documented with several RNA viruses and mutagenic base and nucleoside analogues, which are licensed for clinical use (reviewed in Ref. [81]). These include, the pioneering works with HIV-1 [80,117], PV [118], FMDV [119,120], LCMV [121,122], and HCV [123,124] in cell cultures. Several base and nucleoside purine and pyrimidine analogues have contributed to the research and development of lethal mutagenesis of RNA viruses. Some of them have been in clinical use for years or decades, notably ribavirin and favipiravir. They have been administered either as Food and Drug Administration-approved agents for some specific diseases, or as off-label use for other diseases. Ribavirin has been extensively applied to treat respiratory syncytial virus infections of infants, and as part of combination therapies for chronic hepatitis C, or as off-label treatment of arenavirus-associated hemorrhagic disease. Favipiravir is used as anti-viral influenza in Japan. Regarding their mechanism of activity, the base or nucleoside analogues (or their pro-drugs) are intracellularly converted into their corresponding nucleoside-triphosphate forms, which are the active mutagenic agents during viral RNA synthesis. They are incorporated by the viral RNA-dependent RNA-polymerases into nascent RNA in competition with the standard purine or pyrimidine nucleotides, giving rise to mutated viral RNA progeny. Indications that lethal mutagenesis is operating include an increase of viral mutation frequency, decrease of specific infectivity and invariance of the consensus sequence of the population. Viral extinction can be achieved either by lethal defection ─which entails moderate mutagenic activity and interference by a class of defective genomes termed defectors─ or by overt lethality when the excess mutations is incompatible with virus infectivity [125]. Analogues whose antiviral activity is exerted partly by lethal mutagenesis have proven effective in producing significant viral load

reductions or the extinction of at least fifteen RNA viruses in cell culture or in vivo [concept and specific examples reviewed in Ref. [81].

More recently, the pyrimidine analogue molnupiravir has been used as an anti-COVID-19 agent. It adds to the delayed chain terminator remdesivir (and its combinations with fluoexetine, itraconazole or other drugs) as first-line treatments against COVID-19. Several potential problems have been identified associated with the use of mutagenic analogues in general and with molnupiravir for COVID-19 in particular. They include side-effects on the host cells and organisms by the analogues or their catabolic products, selection of viral mutants resistant to the analogues, or selection of additional mutants when the treatment is only partially effective and analogue-produced mutants retain sufficient replicative and epidemiological fitness. The latter scenario has been suggested for molnupiravir treatments as the possible origin of SARS-CoV-2 mutants when the drug is administered to immunocompromised patients that exhibited a debilitated immune response, thus allowing large replicating viral loads [126,127]. Virological, biochemical and structural studies suggest that lethal mutagenesis is, at least in part, the mechanism of some antiviral agents currently used to treat COVID-19 [115,128–130] .

Synergistic antiviral combinations that include one or two lethal mutagens (a strategy termed synergistic lethal mutagenesis) constitute a promising avenue for the treatment of COVID-19, extensible to other emergent RNA viral infections [115]. Additionally, lethal defection, characterized by the extinction of viral populations due to the emergence of DVGs, has been observed for LCMV in cell cultures [125]. For this virus, two extinction pathways have been identified: one at high mutagen concentrations, resulting in the complete loss of infectivity and replication ability of the quasispecies, and another at lower mutagen concentrations, where replication persists while the infective class becomes extinct due to the presence of defectors.

For a long time, the aforementioned cb/sb mutants have been known for their negative impact on virus accumulation, as they retain the replication signals of both polarities and have a replication advantage owed to their very short genomes. In recent years, the antiviral efficiency of synthetic cb/sb, known as therapeutic interfering particles (TIPs), has been proven against SARS-CoV-2 in animal models. These antiviral effects have involved a decrease in virulence and milder edema symptomatologies [131] as well as a decreased viral transmission [132] in hamsters. Thus, the continuous expansion of theoretical quasispecies concepts, and their scrutiny with viral experimental systems, now reinforced with the new tools of ultra-deep sequencing, is paving the way towards innovations for the control of highly heterogeneous and dynamic cellular pathogens.

References

1. Eigen, M. (1971). Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, *58*, 465-523.
2. Eigen, M., & Schuster, P. (1979). *The Hypercycle: A Principle of Natural Self-Organization*. Springer Berlin Heidelberg.
3. Swetina, J., & Schuster, P. (1982). Self-replication with errors: a model for polynucleotide replication. *Biophys. Chem.*, *16*, 329-345.
4. Eigen, M., McCaskill, P. J., & Schuster, P. (1989). *Molecular quasi-species* (Advances in Chemical Physics. John Wiley & Sons, Inc. ed., Vol. LXXV). I. Prigogine, Stuart A. Rice.
5. Domingo, E., Holland, J. J., & Webster, R. G. (Eds.). (1999). Origin and Evolution of Viruses. Elsevier Science.
6. Domingo, E. (2001). Quasispecies and RNA Virus Evolution: Principles and Consequences (E. Domingo, Ed.). Landes Bioscience.
7. Mas, A., Ulloa, E., Bruguera, M., Furčić, V., & et al. (2004). Hepatitis C virus population analysis of a single-source nosocomial outbreak reveals an inverse correlation between viral load and quasispecies complexity. *J. Gen. Virol.*, 85(12), 3619-3626.
8. Biebricher, C., & Eigen, M. (2005). What is a quasispecies? *Curr. Top. Microbiol. Immunol.*, 299, 1-31.
9. Elena, S. F., Solé, R., & Saranyés, J. (2010). Simple genomes, complex interactions: Epistasis in RNA virus. *Chaos*, 20, 026106.
10. Domingo, E., García-Crespo, C., & Perales, C. (2021). Historical perspective on the discovery of the quasispecies concept. *Ann. Rev. Virol.*, 8, 51-72.
11. Solé, R. V. (2003). Phase transitions in unstable cancer cell populations. *Eur. Phys. J.*, 35, 117-123.
12. Solé, R. V., & Deisboek, T. (2004). An error catastrophe in cancer? *J. Theor. Biol.*, 228, 47-54.
13. Brumera, Y., Michor, F., & Shakhnovich, E. I. (2006). Genetic instability and the quasispecies model. *J. Theor. Biol.*, 241, 216-222.
14. Li, J., Browning, S., Mahal, S., Oelschlegel, A., & et al. (2010). Darwinian evolution of prions in cell culture. *Science*, 327, 869-872.
15. Domingo, E., Sabo, D., Taniguchi, T., & Weissman, C. (1978). Nucleotide sequence heterogeneity of an RNA phage population. *Cell* 13(4), 735-744.
16. Schuster, P., & Stadler, P. (2023). Viral Fitness and Evolution. Current Topics in Microbiology and Immunology (E. Domingo, P. Schuster, S. F. Elena, & C. Perales, Eds.; Vol. 439). Springer International Publishing.
17. Van Cleve, J., & Weissman, D. B. (2015). Measuring ruggedness in fitness landscapes. *Proc Natl Acad Sci USA*, 112(24), 7345-7346.
18. Domingo, E., Soria, M., Gallego, I., de Ávila, A., & et al. (2020). A new implication of quasispecies dynamics: Broad virus diversification in absence of external perturbations. *Infect Genet Evol*, 82, 104278.
19. Domingo, E., Schuster, P., Elena, S. F., & Perales, C. (Eds.). (2023). Viral Fitness and Evolution: Population Dynamics and Adaptive Mechanisms. Springer International Publishing.
20. Elena, S., & Moya, A. (1999). Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. *J. Evol. Biol.*, 12(6), 1078-1088.
21. de la Peña, M., Elena, S. F., & Moya, A. (2000). Effect of deleterious mutation-accumulation on the fitness of RNA bacteriophage MS2. *Evolution* 54(2), 686-691.
22. Sanjuán, R., Moya, A., & Elena, S. F. (2004). The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. USA*, 101, 8396-8401.
23. Carrasco, P., de la Iglesia, F., & Elena, S. F. (2007). Distribution of fitness and virulence effects caused by single-nucleotide substitutions in Tobacco etch virus. *J. Virol.*, 81, 979-984.
24. Fernández, G., Clotet, B., & Martínez, M. A. (2007). Fitness landscape of human immunodeficiency virus type 1 protease quasispecies. *J. Virol.*, 81(5), 1485-2496.
25. Sanjuán, R. (2010). Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. Philos. Trans. *R. Soc. Lond. B Biol. Sci.*, 365(1548), 1975-1982.
26. Kauffman, S. A., & Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *J. Theor. Biol.*, 128(1), 11-45.
27. Schuster, P. (2016). Quasispecies on fitness landscapes. In Quasispecies: From Theory to Experimental Systems (pp. 61-120). E. Domingo, P. Schuster.
28. Lalić, J., & Elena, S. F. (2015). The impact of high-order epistasis in the within-host fitness of a positive-sense plant RNA virus. *J Evol Biol*, 28, 2236-2247.
29. Cervera, H., Lalić, J., & Elena, S. F. (2016). Effect of host species on the topography of the fitness landscape for a plant RNA virus. *J Virol*, 90, 10160-10169.

30. Cervera, H., Lalić, J., & Elena, S. F. (2016). Efficient escape from local optima in a highly rugged fitness landscape by evolving RNA virus populations. *Proc R Soc B*, 283, 20160984.

31. Domingo, E., Martínez-González, B., García-Crespo, C., Somovilla, P., & et al. (2023). Puzzles, challenges, and information reservoir of SARS-CoV-2 quasispecies. *J Virol*, 97(12).

32. Bull, J. J., Meyers, L. A., & Lachmann, M. (2005). Quasispecies made simple. *PLoS Comp. Biol.*, 1(6), e61.

33. Solé, R. V., Sardanyés, J., Díez, J., & Mas, A. (2006). Information catastrophe in RNA viruses through replication thresholds. *J. Theor. Biol.*, 240, 353-359.

34. Biebricher, C., & Eigen, M. (2005). The error threshold. *Virus. Res.*, 107(2), 117-127.

35. Sardanyés, J., & Elena, S. F. (2010). Error threshold in RNA quasispecies models with complementation. *J. Theor. Biol.*, 265, 278-286.

36. Solé, R. V., Sardanyés, J., & Elena, S. F. (2021). Phase transitions in virology. *Rep. Progr. Phys.*, 84, 115901.

37. Nowak, M., & Schuster, P. (1989). Error thresholds of replication in finite populations mutation frequencies and the onset of Muller's ratchet. *J. Theor. Biol.*, 137, 375-395.

38. Sardanyés, J., Elena, S. F., & Solé, R. V. (2008). Simple quasispecies models for the survival-of-the-flattest effect: the role of space. *J. Theor. Biol.*, 250, 560-568.

39. Ben-Ari, I., & Schinazi, R. (2016). A stochastic model for the evolution of a quasispecies. *J. Stat. Phys.*, 162, 415-425.

40. Altemeyer, S., & McCaskill, J. (2001). Error threshold for spatially resolved evolution in the quasispecies model. *Phys. Rev. Lett.*, 86, 5819.

41. Aguirre, J., & Manrubia, S. (2008). Effect of spatial competition on the diversity of a quasispecies. *Phys. Rev. Lett.*, 100, 038106.

42. Sardanyés, J., & Elena, S. F. (2011). Quasispecies spatial models for RNA viruses with different replication modes and infection strategies. *PLoS ONE*, 6, e24884

43. Jacobi, M. N., & Nordahl, M. (2006). Quasispecies and recombination. *Theor. Pop. Biol.*, 70(4), 479-485.

44. Saakian, D. B., Koh, J. M., & Cheong, K. H. (2019). Approximate perturbative solutions of quasispecies model with recombination. *Phys. Rev. E*, 99, 062407.

45. Denhardt, D., & Silver, R. B. (1966). An analysis of the clone size distribution of φX174 mutants and recombinants. *Virology*, 30, 10-19.

46. Chao, L., Rang, C. U., & Wong, L. E. (2002). Distribution of spontaneous mutants and inferences about the replication mode of the RNA bacteriophage φ6. *J. Virol.*, 76, 3276-3281.

47. García-Villada, L., & Drake, J. W. (2012). The three faces of riboviral spontaneous mutation: spectrum, mode of genome replication, and mutation rate. *PLoS Genet.*, 8(7), e1002832.

48. Martínez, F., Sardanyés, J., Elena, S. F., & Daròs, J. A. (2011). Dynamics of a plant RNA virus intracellular accumulation: stamping machine vs. geometric replication. *Genetics* 188(3), 637-646.

49. Schulte, M. B., Draghi, J. A., Plotkin, J. B., & Andino, R. (2015). Experimentally guided models reveal replication principles that shape the mutation distribution of RNA viruses. *eLife* 4, e03753.

50. Sardanyés, J., Solé, R., & Elena, S. F. (2009). Replication mode and landscape topology differentially affect RNA virus mutational load and robustness. J. *Virol.*, 83(23), 12579-12589.

51. Schuster, P., & Swetina, J. (1988). Stationary mutant distributions and evolutionary optimization. *Bull. Math. Biol.*, 50, 635-660.

52. van Nimwegen, E., Crutchfield, J. P., & Huynen, M. (1999). Neutral evolution of mutational robustness. *Proc. Natl. Acad. Sci. USA*, 96, 9176-9720.

53. Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., & et al. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature* 412(19), 331-333.

54. Codoñer, F. M., Daròs, J. A., Solé, R. V., & Elena, S. F. (2006). The fittest versus the flattest: experimental confirmation of the quasispecies effect with subviral pathogens. *PLoS Pathog.*, 2(12), e136.

55. Sanjuán, R., Cuevas, J. M., Holmes, E. C., & Moya, A. (2007). Selection for robustness in mutagenized RNA viruses. *PLoS Genet.*, 3(6), e93.

56. Gregori, J., Colomer-Castell, S., Ibañez-Lligoña, M., García-Cehid, D., & et al. (2024). In-host flat-like quasispecies: characterization methods and clinical implications. *Microroganisms*, 12(5), 1011.

57. Saakian, D. B., & Hu, C.-K. (2006). Exact solution of the Eigen model with general fitness functions and degradation rates. *Proc. Natl. Acad. Sci. USA*, 103(13), 4935-4939.

58. Wilke, C. O., & Ronnewinkel., C. (2001). Dynamic fitness landscapes: expansions for small mutation rates. *Phys. A*, 290, 475-490.

59. Wilke, C. O., Ronnewinkel, C., & Martinetz, T. (2001). Dynamic fitness landscapes in molecular evolution. *Phys. Rep.*, 349, 395-446.

60. Campagnola, G., Govindarajan, V., Pelletier,, A., Canard, B., & et al. (2022). The SARS-CoV nsp12 polymerase active site is tuned for large-genome replication. *J. Virol.*, 96, e0067122.

61. Yin, X., Stapon, A., Bouda, E., & Garcia-Diaz, M. (2023). Fidelity of ribonucleotide incorporation by the SARS-CoV-2 replication complex. *J. Mol. Biol.*, 435(5), 167973.

62. Mackey, M. C., & Glass, L. (1977). Oscillation and chaos in physiological control systems. *Science*, *197*(4300), 287-289.

63. Kopek, B.G., Perkins, G., Miller, D.J., Ellisman, M.H., & et al. (2007). Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol.*, *5*, e220.

64. Harak, C., & Lohmann, V. (2015). Ultrastructure of the replication sites of positive-strand RNA viruses. *J. Virol.*, *419*, 418-433.

65. Honjo, M., Emura, N., Kawagoe, T., Sujisaka, J., & et al. (2020). Seasonality of interactions between a plant virus and its host during persistent infection in a natural environment. *ISME J.*, *14*, 506-518.

66. Zandi, M., Shafaati, M., Shapshak, P., & Reza Hashemnia, S. M. (2023). Monkeypox virus replication underlying circadian rhythm networks. *J. Neurovirol.*, *29*(1), 1-7.

67. Perales, C. (2020). Quasispecies dynamics and clinical significance of hepatitis C virus (HCV) antiviral resistance. *Int. J. Antimicrob. Agents*, *56*(1), 105562.

68. Martínez-González, B., Vázquez-Sirvent, L., Soria, M.E., Mínguez, P., & et al. (2022). Vaccine breakthrough infections with SARS-CoV-2 Alpha mirror mutations in Delta Plus, Iota, and Omicron. *J. Clin. Inv.*, *132*(9), e157700.

69. Delgado, S., Somovilla, P., Ferrer-Orta, C., Martínez-González, B., & et al. (2024). Incipient functional SARS-CoV-2 diversification identified through neural network haplotype maps. *Proc. Natl. Acad. Sci.*, *121*(10), e2317851121.

70. Martínez-González, B., Vázquez-Sirvent, I., Soria, M.E., Mínguez, P., & et al. (2022). SARS-CoV-2 mutant spectra at different depth levels reveal an overwhelming abundance of low frequency mutations. *Pathogens*, *11*, 662.

71. Vignuzzi, M., & López, C.B. (2019). Defective viral genomes are key drivers of the virus–host interaction. *Nat. Microbiol.*, *4*, 1075-1087.

72. González-Aparicio, L., & López, C. (2024). Selection of nonstandard viral genomes during the evolution of RNA viruses: a virus survival strategy of a pesky inconvenience? *Adv. Virus Res.*, *119*, 39-61.

73. Olmo-Ulceda, M. J., Muñoz-Sánchez, J. C., Lasso-Giraldo, W., Arnau, V., & et al. (2022). DVGfinder: a metasearch engine for identifying defective viral genomes in RNA-Seq data. *Viruses*, *14*(5), 1114.

74. Hillung, J., Olmo-Uceda, M. J., Muñoz-Sánchez, J. C., & Elena, S. F. (2024). Accumulation dynamics of defective genomes during experimental evolution of two betacoronaviruses. *Viruses*, *16*(4), 644.

75. Rangel, M. A., Dolan, P. T., Taguwa, S., & et al. (2023). High-resolution mapping reveals the mechanism and contribution of genome insertions and deletions to RNA virus evolution. *Proc. Natl. Acad. Sci. USA*, *120*(31), e2304667120.

76. Martin, M. A., Berg, N., & Koelle, k. (2024). Influenza A genomic diversity during human infections underscores the strength of genetic drift and the existence of tight transmission bottlenecks. *Virus Evolution*, *10*(1), veae042.

77. Zhou, T., Gilliam, N. J., Li, S., Spandau, S., & et al. (2023). Generation and functional analysis of defective viral genomes during SARS-CoV-2 infection. *mBio, 14*(3), e0025023.

78. Campos, C., Colomer-Castell, S., Garcia-Cehic, D., Gregori, J., & et al. (2022). The frequency of defective genomes in Omicron differs from that of the Alpha, Beta and Delta variants. *Sci. Rep.*, *12*, 22571.

79. Domingo-Calap, P., Cuevas, J. M., & Sanjuán, R. (2009). The fitness effects of random mutations in single-stranded DNA and RNA bacteriophages. *PLoS Genet.*, *5*(11), e1000742.

80. Loeb, L., Essigmann, J., Kazazi, F., Zhang, J., & et al. (1999). Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc. Natl. Acad. Sci. USA, 96*(4), 1492-1497.

81. Perales, C., Gallego, I., de Ávila, A.I., Soria, M.E., & et al. (2019). The increasing impact of lethal mutagenesis of viruses. *Future Med. Chem.*, *11*(13), 1645-1657.

82. Iwasa, Y., Michor, F., & Nowak, M. A. (2003). Evolutionary dynamics of escape from biomedical intervention. *Proc Biol Sci*, *270*(1533), 2573-8.

83. Beerenwinkel, N., Eriksson, N., & Sturmfels, B. (2006). Evolution on distributed lattices. *J Theor Biol, 242*(2), 409-420.

84. Traulsen, A., Iwasa, Y., & Nowak, M. A. (2007). The fastest evolutionary trajectory. *J Theor Biol*, *249*(3), 617-623.

85. Perales, C., Agudo, R., Tejero, H., Manrubia, S. C., & et al. (2009). Potential benefits of sequential inhibitor-mutagen treatments of RNA virus infections. *PLoS Path*, *5*(11), e1000658.

86. von Kleist, M., Menz, S., Stocker, H., Aratesh, K., & et al. (2011). HIV quasispecies dynamics during pro-active treatment switching: Impact on multi-drug resistance and resistance archiving in latent reservoirs. *PLoS ONE*, *6*(3), e18204.

87. Sardanyés, J., & Solé, R. V. (2013). Red Queen coevolution on fitness landscapes. In H. Richter & A. Engelbrecht (Eds.), *Recent Advances in the Theory and Application of Fitness Landscapes* (pp. 301-338). Springer Berlin Heidelberg.

88. Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Phys. Rev. Lett.*, *86*, 3200.

89. Zhang, X., Ruan, Z., Zheng, M., & Zhou, J. (2022). Epidemic spreading under mutually independent intra- and inter-host pathogen evolution. *Nature Communications*, *13*(6218), 1-13.

90. Pocock, M. J.O., Evans, D. M., & Memmott, J. (2012). The robustness and restoration of a network of ecological networks. *Science*, *335*, 973-977.

91. Pilosof, S., Porter, M., Pascual, M., & Kéfi, S. (2017). The multilayer nature of ecological networks. *Nat Ecol. Evol.*, *1*(4), 0101.

92. Cellai, D., López, E., Zhou, J., & Gleeson, J.P. (2013). Percolation in multiplex networks with overlap. *Phys. Rev. E*, *88*, 052811.

93. Boccaletti, S., Bianconi, G., Criado, R., del Genio, C.I., & et al. (2014). The structure and dynamics of multilayer networks. *Phys. Rep.*, *544*(1), 1-122.

94. Dickison, M., Havlin, S., & Stella, H.E. (2012). Epidemics on interconnected networks. *Phys. Rev. E*, *85*, 066109.

95. Mendiola, A.S., Serrano, M., & Boguñá, M. (2012). Epidemic spreading on interconnected networks. *Phys. Rev. E*, *86*, 026106.

96. Zhao, D., Li, L., Peng, H., Luo, Q., & et al. (2014). Multiple routes transmitted epidemics on multiplex networks. *Phys. Lett. A*, *378*(10), 770-776.

97. Stella, M., Andreazzi, C., Selakovic, S., & Goudarzi, A. (2017). Parasite spreading in spatial ecological multiplex networks. *J. Complex Netw.*, *5*, 486-511.

98. Funk, S., & Jansen, V. A. A. (2010). Interacting epidemics on overlay networks. *Phys. Rev. E*, *81*, 036118.

99. Sahneh, F. D., & Scoglio, C. (2014). Competitive epidemic spreading over arbitrary multilayer networks. *Phys. Rev. E*, *89*(6), 062817.

100. Sanz, J., Xia, C.-Y., Meloni, S., & Moreno, Y. (2014). Dynamics of interacting diseases. *Phys. Rev. X*, *4*, 041005.

101. Scatà, M., Di Stefano, A., Liò, P., La Corte, A.(2013) The Impact of Heterogeneity and Awareness in Modeling Epidemic Spreading on Multiplex Networks. *Sci. Rep.*, 6(37105), 1-13.

102. Park, M., Loverdo, C., Schreiber, S. J., & Lloyd-Smith, J. O. (2013). Multiple scales of selection influence the evolutionary emergence of novel pathogen. *Philos.Trans. R. Soc. Lond. B*, *368*(1614), 20120333.

103. Schreiber, S. J., Ke, R., Loverdo, C., Park, M., & et al. (2021). Cross-scale dynamics and the evolutionary emergence of infectious diseases. *Virus Evolution*, *7*(1), veaa105.

104. Morse, S. S. (Ed.). (1993). *Emerging Viruses*. Oxford University Press.

105. Morens, D., & Fauci, A. (2000). Emerging pandemic diseases: how we got to COVID-19. *Cell*, *183*, 837.

106. Smolinski, M. S., Hamburg, M. A., & Lederberg, J. (Eds.). (2003). *Microbial Threats to Health: Emergence, Detection, and Response*. National Academies Press.

107. Solé, R. V. (2000). Signs Of Life: How Complexity Pervades Biology. Basic Books.

108. Solé, R. V., & Elena, S. F. (2018). *Viruses as Complex Adaptive Systems*. Princeton University Press.

109. Domingo, E. (2020). Virus as Populations: Composition, Complexity, Dynamics, and Biological Implications (2nd ed.). Academic Press, Elsevier Inc.

110. Fournelle, D., Mostefai, F., Brunet-Ratnasingham, E., & Poujol, R. (2024). Intra-Host Evolution Analyses in an Immunosuppressed Patient Supports SARS-CoV-2 Viral Reservoir Hypothesis. *Viruses*, *16*(3), 342.

111. Igari, H., Sakao, S., Ishige, T., & Saito, K. (2024). Dynamic diversity of SARS-CoV-2 genetic mutations in a lung transplantation patient with persistent COVID-19. *Nat Comunn*, *15*, 3602.

112. Holland, J. J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., & et al. (1982). Rapid evolution of RNA genomes. *Science*, *215*(4540), 1577-1585.

113. Lederberg, J. (1993). Viruses and humankind. Intracellular symbiosis and evolutionary competition. In S. S. Morse (Ed.), Emerging Viruses (pp. 3-9). Oxford University Press.

114. Organization WH. (2003). *SARS: status of the outbreak and lessons for the immediate future 2003*. WHO. https://www.who.int/csr/media/sars_wha.pdf

115. García-Crespo, C., de Ávila, A.I., gallego, I., Soria, M.E., & et al. (2024). Synergism between remdesivir and ribavirin leads to SARS-CoV-2 extinction in cell culture. *Br. J. Pharmacol.*, 1-19. https://doi.org/10.1111/bph.16344

116. Papastavrou, N., Horning, D. P., & Joyce, G. F. (2024). RNA-catalyzed evolution of catalytic RNA. *Proc. Natl. Acad. Sci. USA*, *121*(11), e2321592121.

117. Dapp, M. J., Patterson, S. E., & Mansky, L. M. (2013). Back to the future: revisiting HIV-1 lethal mutagenesis. Trends Microbiol., 21(2), 56-62.

118. Crotty, S., Cameron, C., & Andino, R. (2001). RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc. Natl. Acad. Sci. USA*, *98*, 6895-6900.

119. Perales, C., Henry, M., Domingo, E., & Wain-Hobson, S. (2011). Lethal mutagenesis of foot-and-mouth disease virus involves shifts in sequence space. *J. Virol.*, *85*, 12227-12240.

120. de Ávila, A., Moreno, E., Perales, C., & Domingo, E. (2017). Favipiravir can evoke lethal mutagenesis and extinction of foot-and-mouth disease virus. *Virus Res.*, *233*, 105-112.

121. Grande-Pérez, A., Sierra, S., Castro, M.G., Domingo, E., & et al. (2002). Molecular indetermination in the transition to error catastrophe: Systematic elimination of lymphocytic choriomeningitis virus through

mutagenesis does not correlate linearly with large increases in mutant spectrum complexity. *Proc. Natl. Acad. Sci. USA, 99*(20), 12938-12943.

122. Ruiz-Jarabo, C. M., Ly, C., Domingo, E., & de la Torre, J. C. (2003). Lethal mutagenesis of the prototypic arenavirus lymphocytic choriomeningitis virus (LCMV). *Virology, 308*(1), 37-47.

123. Ortega-Prieto, A., Sheldon, J., Grande-Pérez, A., & Tejero, H. (2013). Extinction of hepatitis C virus by ribavirin in hepatoma cells involves lethal mutagenesis. *PLoS ONE, 8*, e71039.

124. de Ávila, A., Gallego, I., Sori, M.E., Gregori, J., & et al. (2016). Lethal mutagenesis of hepatitis C virus induced by favipiravir. *PLoS ONE, 11*, e0164691.

125. Grande-Pérez, A., Lázaro, E., Lowenstein, P., Domingo, E., & et al. (2005). Suppression of viral infectivity through lethal defection. *Proc. Natl. Acad. Sci. USA, 102*, 4448-4452.

126. Sanderson, T., Hisner, R., Donovan-Banfield, I., Hartman, H., & et al. (2023). A molnupiravir-associated mutational signature in global SARS-CoV-2 genomes. *Nature, 623*(7987), 594-600.

127. Fountain-Jones, N. M., Vanhaeften, R., Williamson, J., & Maskell, J. (2024). Effect of molnupiravir on SARS-CoV-2 evolution in immunocompromised patients: a retrospective observational study. *The Lancet Microbe, 5*(5), e452-e458.

128. Gordon, C. J., Tchesnokov, E. P., Schinazi, R. F., & Götte, M. (2021). Molnupiravir promotes SARS-CoV-2 mutagenesis via the RNA template. *J. Biol. Chem., 297*(1), 100770.

129. Kabinger, F., Stiller, C., Schmitzová, J., Dienemann, C., & et al. (2021). Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nature Struct. Molec. Biol., 28*, 740-746.

130. Somovilla, P., García-Crespo, C., Martínez-González, B., Soria, M. E., & et al. (2023). Atypical mutational spectrum of SARS-CoV-2 replicating in the presence of ribavirin. *Antimicrob. Agents Chemother., 67*((1)), e0131522.

131. Chaturvedi, S., Vasen, G., Pablo, M., Chen, X., & et al. (2021). Identification of a therapeutic interfering particle—A single-dose SARS-CoV-2 antiviral intervention with a high barrier to resistance. *Cell, 184*(25), 6022-6036.

132. Chaturvedi, S., Beutler, N., Basen, G., Pablo, M., & et al. (2022). A single-administration therapeutic interfering particle reduces SARS-CoV-2 viral shedding and pathogenesis in hamsters. *Proc. Natl. Acad. Sci. USA, 119*(39), e2204624119.

133. Acevedo, A., Brodsky, L., & Andino, R. (2014). Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature, 505*, 686.

134. 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.

135. Bruguera, M., Saiz, J., Franco, S., Giménez-Barcons, M., & et al. (2002). Outbreak of Nosocomial Hepatitis C Virus Infection Resolved by Genetic Analysis of HCV RNA. *J Clin Microbiol, 40*(11), 4363-5366.

136. Duarte, E. A., Novella, I.S., Ledesma, S., Clarke, D.K., & et al. (1994). Subclonal components of consensus fitness in an RNA virus clone. *J. Virol., 68*(7), 4295-4301.

137. Gago, S., Elena, S. F., Flores, R., & Sanjuán, R. (2009). Extremely high mutation rate of a hammerhead viroid. *Science, 323*(5919), 1308.

138. Holland, J. J., et al, Domingo, E., de la Torre, J., & Steinhauer, D.A. (1990). Mutation frequencies at defined single codon sites in vesicular stomatitis virus and poliovirus can be increased only slightly by chemical mutagenesis. *J. Virol., 64*, 3960-3962.

139. Mansky, L., & Temin, H. (1995). Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase. *J. Virol., 69*(8), 5087-5094.

140. Martínez-González, B., Soria, M.E., Mínguez, P., Lorenzo-Redondo, R., & et al. (2024). SARS-CoV-2 mutant spectra as variant of concern nurseries: endless variation? *Front. Microb., 15*, 1358258.

141. Novella, I., Duarte, E.A., Elena, S. F., Moya, A., & et al. (1992). Exponential increases of RNA virus fitness during large population transmissions. *Proc. Natl. Acad. Sci. USA, 92*, 5841-5844.ç

142. Ruiz-Jarabo, C. M., Arias, A., Baranowski, E., Escarmís, C., & et al. (2000). Memory in viral quasispecies. *J. Virol. 74, 8*, 3543-3547.