



## Article

# Viral Quasispecies Inference from Single Observations—Mutagens as Accelerators of Quasispecies Evolution

Josep Gregori <sup>1,\*</sup> , Miquel Salicrú <sup>2</sup> , Marta Ibáñez-Lligoña <sup>1,3,4</sup> , Sergi Colomer-Castell <sup>1,3,5</sup> , Carolina Campos <sup>1,3,5</sup> , Alvaro González-Camuesco <sup>1,3</sup> and Josep Quer <sup>1,3,4,5,\*</sup>

- <sup>1</sup> Liver Diseases-Viral Hepatitis, Liver Unit, Vall d'Hebron Institute of Research (VHIR), Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain; marta.ibanez@vhir.org (M.I.-L.); sergi.colomer@vhir.org (S.C.-C.); carolina.campos@vhir.org (C.C.); alvaro.gonzalez@vhir.org (A.G.-C.)
- <sup>2</sup> Statistics Department, Biology Faculty, University of Barcelona, 08028 Barcelona, Spain; msalicru@ub.edu
- <sup>3</sup> Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Instituto de Salud Carlos III, Av. Monforte de Lemos, 3-5, 28029 Madrid, Spain
- <sup>4</sup> Medicine Department, Universitat Autònoma de Barcelona (UAB), Campus de la UAB, Plaça Cívica, 08193 Bellaterra, Spain
- <sup>5</sup> Biochemistry and Molecular Biology Department, Universitat Autònoma de Barcelona (UAB), Campus de la UAB, Plaça Cívica, 08193 Bellaterra, Spain
- \* Correspondence: josep.gregori@gmail.com (J.G.); josep.quer@vhir.org (J.Q.)

## Abstract

RNA virus populations exist as quasispecies—complex, dynamic clouds of closely related but genetically diverse variants generated by high mutation rates during replication. Assessing quasispecies structure and diversity is crucial for understanding viral evolution, adaptation, and response to antiviral treatments. However, comparing single quasispecies observations from individual biosamples, especially at different infection or treatment time points, presents statistical challenges. Traditional inferential tests are inapplicable due to the lack of replicate observations, and resampling-based approaches such as the bootstrap and jackknife are limited by biases and non-independence, particularly for diversity indices sensitive to rare haplotypes. In this study, we address these limitations by applying the delta method to derive analytical variances for a set of quasispecies structure indicators specifically designed to assess the quasispecies maturation state. We demonstrate the utility of this approach using high-depth next-generation sequencing data from hepatitis C virus (HCV) quasispecies evolving in vitro under various conditions, including free evolution and exposure to antiviral or mutagenic treatments. Our results reveal that with highly fit HCV quasispecies, sofosbuvir inhibits quasispecies genetic diversity, while mutagenic treatments accelerate maturation, compared to untreated controls. We emphasize the interpretation of results through absolute differences, log-fold changes, and standardized effect sizes, moving beyond mere statistical significance. This framework enables robust, quantitative comparisons of quasispecies diversity from single observations, providing valuable insights into viral adaptation and treatment response. The R code and session info with required libraries and versions is provided in the supplementary material.

**Keywords:** inference; quasispecies; single-observation; RNA-virus; mutagens; inhibitors



Academic Editor: Flor Helene Pujol

Received: 28 July 2025

Revised: 25 August 2025

Accepted: 28 August 2025

Published: 30 August 2025

**Citation:** Gregori, J.; Salicrú, M.; Ibáñez-Lligoña, M.; Colomer-Castell, S.; Campos, C.; González-Camuesco, A.; Quer, J. Viral Quasispecies Inference from Single Observations—Mutagens as Accelerators of Quasispecies Evolution. *Microorganisms* **2025**, *13*, 2029. <https://doi.org/10.3390/microorganisms13092029>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

A quasispecies is the complex and dynamic population of closely related but genetically diverse viral variants that arises during RNA virus infection. This population structure

arises due to the high mutation rates of RNA viruses, which can introduce multiple errors during genome replication. The quasispecies consists of a mutant spectrum or cloud of viral genomes that are continuously generated and subject to selection within the host [1–3]. Quasispecies maturation is understood as an enrichment in genetic diversity caused by replication errors. This process facilitates the selection and proliferation of alternative genetic variants expressing functional phenotypes, resulting in enhanced quasispecies fitness, increased resilience and reduced susceptibility to antiviral treatments [4–7].

A quasispecies observation consists in sequencing to high depth a quasispecies sample taken from a biosample infected with a RNA virus, and it is represented by a set of genomes (haplotypes) and observed frequencies, in the form of NGS sequencing read counts. This observation constitutes a sample of size  $n$ , with  $n$  the number of final reads.

In the study of quasispecies structure and diversity, we are often confronted with a situation where we need to compare single quasispecies observations from single biosamples, possibly taken at two different infection or treatment time points. This situation presents statistic challenges, because traditional inferential tests are not applicable when comparing diversity indices or structure indicators of single quasispecies observations. A useful test must be based on estimates of expected diversity and its variability for each of the quasispecies to be compared. Analytical approximations based on the Delta method are available for any index or indicator expressed as a function of multinomial proportions [8–10]. Besides, we know that some diversity indices, which may be of interest, are sensitive to sample size differences, which means that unbalanced samples will require normalization. We focus on diversity indices which may be calculated directly from the vector of haplotype NGS read counts.

Alternative methods rely on tools based on resampling to generate empirical distributions from which estimate standard errors, confidence intervals, and  $p$ -values. A resampling method commonly used to generate such distributions to estimate variances is the bootstrap [11–13], where at each resampling cycle, we sample with replacement a number of reads equal to the sample size. At each bootstrap cycle diversity indices are computed, and at the end the estimated values and corresponding variances result from the mean and variance of the observed values in the  $B$  bootstrap cycles. However we know that the bootstrap suffers from the 0.632 limitation, which implies that in each resampling cycle with replacement at the full sample size, at most 63.2% of the reads in the sample are observed, meaning that 36.8% of the reads are replicates of already observed reads, and likewise that 36.8% of the reads in the quasispecies sample remain unobserved [14]. This limitation results in biases in the estimation of diversity indices sensitive to low and very low haplotype frequencies. Particularly affected are the fraction of singletons and rare haplotypes, the total number of haplotypes and indices such as Shannon entropy or the Hill number at  $q = 1$  [14]. Note that any approach based on parametric multinomial resampling is submitted to the same limitations; resampling from a multinomial given its parameters and a sample size, is equivalent to resampling with replacement.

An alternative to the bootstrap is the jackknife [11]. While similar to the bootstrap in that it involves resampling, the key difference is that the jackknife samples without replacement, which aligns with our requirements. However, it is well known that the traditional delete-1 jackknife estimator performs poorly with non-smooth estimators, those sensitive to one or a few values in the dataset, such as the median or other empirical quantiles. This makes it unsuitable for computing confidence intervals. It has been shown that the delete- $d$  Jackknife [11], where  $d/n \rightarrow \delta \in (0, 1)$  asymptotically (rather than the overly restrictive  $\sqrt{n} < d < (n - 1)$ ), is a consistent estimator [11,15] and overcomes most of the limitations in the delete-1 approach. This method resembles the delete-1 estimator but incorporates a normalization constant accounting for the number of distinct subsets with  $d$

items removed  $\binom{n}{d}$ . However, computational complexity increases rapidly with larger  $d$ , especially for large  $n$ . To address this, the estimator is often evaluated using a subset of all possible  $\binom{n}{d}$  sets [11,16,17]. A key limitation is that estimated variances depend on  $d$ , the number of reads removed in each iteration. Furthermore, the jackknife values generated under this scheme are not independent and identically distributed (i.i.d.), as resampled subsets overlap. Having i.i.d. values is a fundamental requirement for obtaining  $p$ -values or confidence intervals from empirically distributed values.

We find ourselves restricted to the use of the delta method, as no resampling scheme appears to be free of significant limitations in our context.

We present under Methods a set of quasispecies structure indicators specifically devised to assess the maturation state of a quasispecies [7] along with their analytical variances obtained using the delta method [18–20]. The Results section illustrates the application of the proposed method to test data of HCV quasispecies evolving in vitro in a non-coevolving environment, either freely or under antiviral treatments [21–24]. Treatments are administered to high-fitness quasispecies evolved over 100 passages in a non-coevolving environment, which show reduced response to antiviral treatments. These illustrative results demonstrate sofosbuvir’s inhibitory effects on quasispecies genetic structure and the accelerated quasispecies maturation caused by mutagen treatments in monotherapy, compared to treatment-free evolution in a non-coevolving environment. We focus on interpreting the results by integrating absolute differences, log-fold changes, and standardized effect sizes, moving beyond statistical significance alone—which is almost guaranteed by the large sample sizes. It must be emphasized, however, that the cell-culture data presented here serve only as an example of the method’s applicability. While the observed trends are consistent with findings from clinical datasets [4–7], these data alone do not support sufficiently robust conclusions, and further studies—including biological replicates of cell-culture experiments—will be needed. The directions of the effects are clear, but their magnitudes may vary across independent studies.

## 2. Materials and Methods

### 2.1. Quasispecies Maturity Indicators

Quasispecies maturation reflects a global trend of within-host quasispecies evolution toward enriched genetic diversity and phenotypic redundancy, driven by continuous emergence and selection of new functional variants with fitness advantages [7]. Table 1 shows the set of selected quasispecies structure indicators to account for the evolutionary state of a quasispecies in its path from a founding quasispecies,  $A$ , dominated by a highly abundant master genome and a peaked fitness landscape, towards its hypothetical endpoint,  $Z$ , in the form of a flattened fitness landscape, with a high number of genomes at low and equivalent abundances, in dynamic equilibrium, with no dominance [5–7,25].

**Table 1.** Selected maturity indicators. Description and expected values in the two limiting states [25].

Indicator	Description	State A	State Z
TopN	Fraction of reads for top N hpl.	1.00	0.00
Master	Dominant haplotype frequency	1.00	0.00
Rare1	Fraction of reads for hpl $\leq 1\%$	0.00	1.00
Rare2	Fraction of reads for hpl $\leq 0.1\%$	0.00	1.00
$RLE_1$	Relative logarithmic evenness at $q = 1$	0.00	1.00
$RLE_2$	Relative logarithmic evenness at $q = 2$	0.00	1.00
$RLE_\infty$	Relative logarithmic evenness at $q = \infty$	0.00	1.00
$R_k$	Evenness among top k haplotypes	0.00	1.00

The zeroes for TopN and Master, in State Z, should be interpreted as an infinitesimal frequency  $\epsilon$ , that is  $\epsilon > 0$  and  $\epsilon < 1/n$  for a big positive integer  $n$ , rather than strictly zero. In biochemical terms this  $\epsilon$  may be interpreted as one to a few or several molecules, in the context of a high viral titer. Similarly, the zeroes for Rare1 and Rare2, in state A, should be interpreted as a result of several variants other than the master haplotype, at infinitesimal frequencies,  $\epsilon$ . However, we represent  $\epsilon$  formally as 0.

A global quasispecies maturity score for a given quasispecies,  $Q_s$ , may be computed as the distance of this quasispecies from State A,  $d_A$ , in the multidimensional space represented by this set of indicators.

$$d_A(Q_s) = \sqrt{\sum_i [(f_i(Q_s) - f_i(A))^2]} \quad (1)$$

where  $f_i$  is the set of maturity indicators. This score may be normalized to the range 0–1 as  $d_N(Q_s) = d_A(Q_s)/d_A(Z)$ . Whereas each indicator refers to a given aspect of quasispecies structure,  $d_N$  represents the combined effect of all changes in pushing the quasispecies towards its purest origin, A, or its fate, Z [7].

## 2.2. Delta Method

The delta method [18,19] provides a statistical foundation for justifying normality in tests involving functions of random variables by leveraging the asymptotic normality of the parameters estimators. When an estimator is asymptotically normal, the delta method allows us to approximate the distribution of smooth (i.e. differentiable) functions of that estimator as also normal. The proposed quasispecies structure indicators are based on the distribution of haplotypes, which follows a multinomial distribution with as many categories and parameters as there are haplotypes. Multinomial proportions estimated from a sample exhibit asymptotically normal behavior,

$$\sqrt{n}(\hat{P} - P) \xrightarrow{d} \mathcal{N}(0, \Sigma_P), \quad \Sigma_P = (\sigma_{ij}(P)) = (\delta_{ij}p_i - p_i p_j)$$

providing the basis for applying the delta method to functions of estimated proportions  $\hat{P} : \{\hat{p}_i; i = 1 \dots H\}$ , and proportion aggregates, and for conducting tests based on the normal distribution using these indicators. In what follows, let  $\hat{P}$  denote the vector of estimated haplotype frequencies, sorted in decreasing order of frequency without loss of generality, so that  $\hat{p}_1$  represents the estimated frequency of the master haplotype.

Brevly, if an estimator  $\hat{\Theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_H)$  satisfies

$$\sqrt{n}(\hat{\Theta} - \Theta) \xrightarrow{d} \mathcal{N}(0, \Sigma_{\Theta}), \quad \Sigma_{\Theta} = (\sigma_{ij}(\Theta))$$

then for a differentiable function  $g$ , the delta method ensures:

$$\sqrt{n}(g(\hat{\Theta}) - g(\Theta)) \xrightarrow{d} \mathcal{N}(0, g'(\Theta)^T \Sigma_{\Theta} g'(\Theta))$$

provided that  $g'(\Theta)^T \Sigma_{\Theta} g'(\Theta) > 0$ , being

$$g'(\Theta)^T = \left( \frac{\partial g(\Theta)}{\partial \theta_1}, \dots, \frac{\partial g(\Theta)}{\partial \theta_H} \right)$$

This allows normality tests to be applied to  $g(\hat{\Theta})$ , even if  $g$  is non linear.

### 2.3. Indicators' Variance and Standard Error by the Delta Method

The method approximates the standard error of  $g(\hat{P})$  of first order Taylor expansion:

$$\begin{aligned} \text{SE}(g(\hat{P}))^2 &= \frac{1}{n} \text{Var}(g(\hat{P})) \approx \frac{1}{n} g'(P)^T \Sigma_P g'(P) \approx \\ &\approx \frac{1}{n} g'(\hat{P})^T \Sigma_{\hat{P}} g'(\hat{P}) \end{aligned}$$

where  $\hat{P}$  is an estimated binomial or multinomial proportion, a vector of multinomial proportions, or an aggregate of them.

The full set of analytically derived equation of the variances for all proposed quasispecies structure indicators using the delta method is provided in the Supplementary Materials [18–20,26].

### 2.4. Large Sample Sizes and $p$ -Values

With sample sizes on the order of 100,000 reads, even minuscule differences in diversity measures between two quasispecies can achieve statistical significance due to the extremely small standard errors. However, statistical significance alone does not guarantee biological or practical relevance. In such cases, 99% confidence intervals help quantify the precision of diversity estimates and delineate the range of plausible values for the true difference between quasispecies. Nevertheless, the primary emphasis should be placed on the magnitude of the observed effect rather than solely on statistical significance. While estimated absolute differences, relative differences, or ratios between indicators of the quasispecies provide useful metrics, these values are not always readily interpretable in biological or clinical contexts. This highlights the need for standardized effect sizes, which allow for meaningful comparison across studies and contexts by expressing the magnitude of differences in a scale-independent manner. Standardized effect sizes facilitate the assessment of whether observed differences are not only statistically significant, but also biologically or clinically meaningful, thereby enhancing the interpretability and generalizability of findings. Effect size measures, like Cohen's  $d$ , can provide more meaningful information about the practical importance of observed differences, and are independent of sample sizes [27–30]. While they may not substitute for biological significance when sufficient knowledge is available, they do offer a generally interpretable scale of magnitudes [27,30].

### 2.5. Effect Size

Effect sizes quantify the magnitude of effects (i.e., strength of a relationship, size of a difference), which are the outcomes of our empirical research. These measures are independent of sample sizes, contrary to  $p$ -values.

In general terms, given a quasispecies indicator  $H$  and the standard error of its estimator,  $\text{SE}^2$ , the effect size for the difference of two quasispecies  $H_A - H_B$ , corresponds to:

$$\begin{aligned} d_H &= \frac{H_A - H_B}{\sqrt{(n_A \cdot \text{SE}_A^2 + n_B \cdot \text{SE}_B^2)/2}} = \\ &= \frac{H_A - H_B}{\sqrt{(\sigma_A^2 + \sigma_B^2)/2}} = \frac{H_A - H_B}{S_p} \end{aligned} \quad (2)$$

It follows from the expression relating the standard error of an estimator and its original standard deviation  $\sigma^2/n = \text{SE}^2$ .

Cohen's  $d$  is a unitless, standardized measure of effect size that quantifies the difference between value estimates in two groups in terms of standard deviation units. When the data is normally distributed, Cohen's  $d$  can be interpreted in probabilistic terms, which offer

a more intuitive interpretation. Two probabilistic metrics can be used: the non-overlap (NOV) of the two distributions and the probability of superiority of one condition over the other. Guidelines for interpreting  $d$  values using common terms of magnitude, as well as their probabilistic equivalents are given in Supplementary Table S1 [30].

### 2.6. $Z$ , $t$ , and Cohen's $d$

With unequal variances the comparison of two estimates is made using the Welch  $t$ -test, where the  $t$ -statistic is defined as:

$$t = \frac{\hat{H}_A - \hat{H}_B}{\sqrt{SE_A^2 + SE_B^2}} = \frac{\hat{H}_A - \hat{H}_B}{\sqrt{s_A^2/n + s_B^2/n}} = d \sqrt{\frac{n}{2}} \quad (3)$$

where  $d$  is the standardized difference based on pooled variance, with  $n_a = n_b = n$ . With a reference size of 100,000 reads, the above equation becomes:

$$t = 224 \times d$$

When the indicator  $H$  corresponds to a proportion, or to an aggregate of frequencies from a multinomial distribution,  $\hat{p}$ , with a reference sample size of  $n = 100,000$  reads, we can use the normal approximation to a proportion:

$$Z = 224 \times \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\bar{p}(1 - \bar{p})}} = 245 \times d$$

where  $\bar{p} = \frac{1}{2}(\hat{p}_A + \hat{p}_B)$ , when  $n_a = n_b = n$ , and  $d$  is the standardized difference based on the pooled variance. The biggest denominator in this equation occurs for  $\bar{p} = 0.5$ , giving 0.5, then:

$$Z \geq 448 \times (\hat{p}_A - \hat{p}_B)$$

The critical  $Z$  value for  $\alpha = 0.0005$  is 3.29. Therefore, a difference in proportion greater than  $3.29/448 = 0.0073$  will be statistically significant at 99.9% level, regardless of the  $\bar{p}$  value.

Note, however, that the probability density function (PDF) of the  $t$ -distribution simplifies to the standard normal PDF as the degrees of freedom tend to infinity. For very high sample sizes,  $n$ , the  $t$ -distribution for the difference of means converges to the standard normal distribution due to the Central Limit Theorem and the reduced uncertainty in estimating the population standard deviation. Therefore, a  $z$ -test is acceptable for all proposed quasispecies indicators in such cases.

### 2.7. Quasispecies Samples, Cells, Viruses, and Drugs

This is a reanalysis of data in previously published works [21,22] that investigates the evolution of HCV viral quasispecies under various conditions in cell culture, including natural passage, direct-acting antiviral (DAA) treatment, and mutagenic agents.

HCV RNA expressed from plasmid Jc1FLAG2(p7-nsGluc2A) (genotype 2a) [31] was transfected into Huh-7 Lunet cells and amplified in Huh-7.5 cells to produce the initial virus population HCV p0 [32]. HCV p0 was subjected to 100 serial passages in Huh-7.5 reporter cells in the absence of any drug, equivalent to  $\approx 300$  days of evolution. The populations at passage 100 (HCV p100) displayed increased replication in Huh-7.5 reported cells [23,33], with increased fitness 2.2-fold relative to HCV p0 [22,33]. The 100 passages are equivalent to about a year of continuous evolution in a non-coevolving environment [23,24] free of external pressures.

HCV p100 was further passaged in the absence of any drug (Ctl), in the presence of favipiravir (FPV, 400  $\mu$ M), ribavirin (RBN, 100  $\mu$ M), and sofosbuvir (SOF, 800 nM), for extra ten passages (30 days). The drug concentrations were chosen for their capacity to extinguish HCV p0 in four to five passages, under the infection conditions used in the experiments [22]. Favipiravir (T-705) (Atomax Chemicals Co., Ltd., Hong Kong, China), ribavirin (Sigma, Hong Kong, China), and sofosbuvir (Selleck Chemicals, Houston, TX, USA) were prepared and used as previously described [21,34,35].

To achieve a comprehensive picture of quasispecies composition, our previous results have underlined the need to perform deep-sequencing studies with very high coverage (over  $1 \times 10^5$  reads per amplicon) and avoid unnecessary abundance-based filter [5–7,25,36]. This is particularly important with mutagenic treatments, where their effect in the short term is mainly observed at the lowest level of haplotype frequencies [5,36].

A NS5B amplicon spanning positions 8552:8869 is deeply sequenced. Sequencing was performed using MiSeq Illumina<sup>®</sup> instruments, in  $2 \times 300$  paired-end mode, obtaining deep coverages ranging from 110 K to 191 K per strand, amplicon and sample (Table 2).

**Table 2.** Reads per sample. Forward strand only.

ID	Reads
p0	123,587
p100	179,382
Ctl	136,432
FPV	191,058
RBV	134,718
SOF	111,299

### 2.8. Implementation

The described statistic methods have been implemented in R functions. R code and session info with R and libraries versions are provided in the Supplementary Material. Given a quasispecies represented by a vector of haplotype frequencies as read counts or relative frequencies, these functions calculate quasispecies indicator values and corresponding variances. These are then used to compare and test pairs of quasispecies. The test provides for each indicator the following information: observed difference, observed ratio, z-statistic, Benjamini-Hochberg multi-test adjusted *p*-values, Conhen's *d*, NOV and PS.

## 3. Results

### 3.1. Maturity Indicator Values on Rarefied Quasispecies

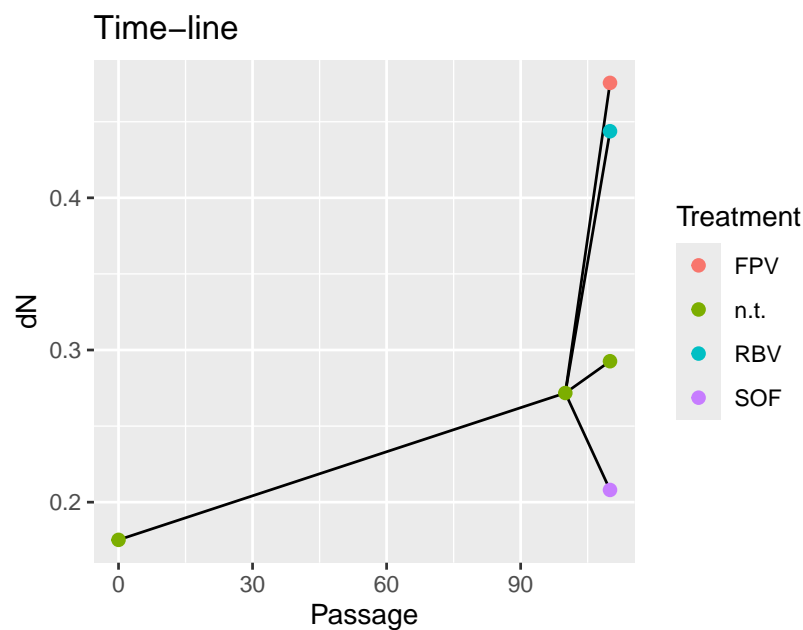
All quasispecies samples are normalized by rarefaction [14] to a reference size of 110,000 reads -the minimum size to include all samples- the median values after 500 cycles of rarefaction are taken for each indicator (Table 3).

The calculated variances from the analytical expressions are displayed in Supplementary Table S2.

Figure 1 provides a summary of the treatment effects on the maturity score *dN*, displaying its evolution along the cell-culture timeline. From p100, ten extra passages without treatment are used as control, to be compared with the results of ten passages, starting at p100, treated with SOF, RBV or FPV in monotherapy. Clearly the treatments with mutagens in ten passages from p100, resulted in accelerated changes on *dN*, greater than those achieved in the 100 passages from p0 of natural evolution in cell culture.

**Table 3.** (A)—Quasispecies maturity indicators. Rarefied values to 110,000 reads. (B)—Quasispecies maturity indicators, continued. Rarefied values to 110,000 reads.

(A)						
ID	nHpl	Master	Top25	Rare1	Rare2	dN
p0	4879.0	0.73728	0.80969	0.23231	0.19061	0.17525
p100	7883.5	0.49492	0.74722	0.32811	0.22594	0.27175
Ctl	8042.0	0.39545	0.75894	0.27743	0.23635	0.29263
SOF	5939.0	0.61619	0.81265	0.22331	0.19370	0.20805
RBV	17,702.5	0.18436	0.55210	0.51516	0.38691	0.44376
FPV	17,138.0	0.12553	0.58100	0.51404	0.37284	0.47553
(B)						
ID	R5	R10	RLE1	RLE2	RLEinf	dN
p0	0.01934	0.02206	0.27130	0.07163	0.03589	0.17525
p100	0.05016	0.05977	0.39069	0.14845	0.07840	0.27175
Ctl	0.08407	0.02905	0.40351	0.18128	0.10317	0.29263
SOF	0.07808	0.02405	0.31369	0.10767	0.05572	0.20805
RBV	0.16893	0.11401	0.57920	0.29061	0.17286	0.44376
FPV	0.42867	0.15571	0.58111	0.31728	0.21286	0.47553

**Figure 1.** Treatments time-line, showing the effects of each treatment on the maturity score  $dN$  over time.

### 3.2. Tests

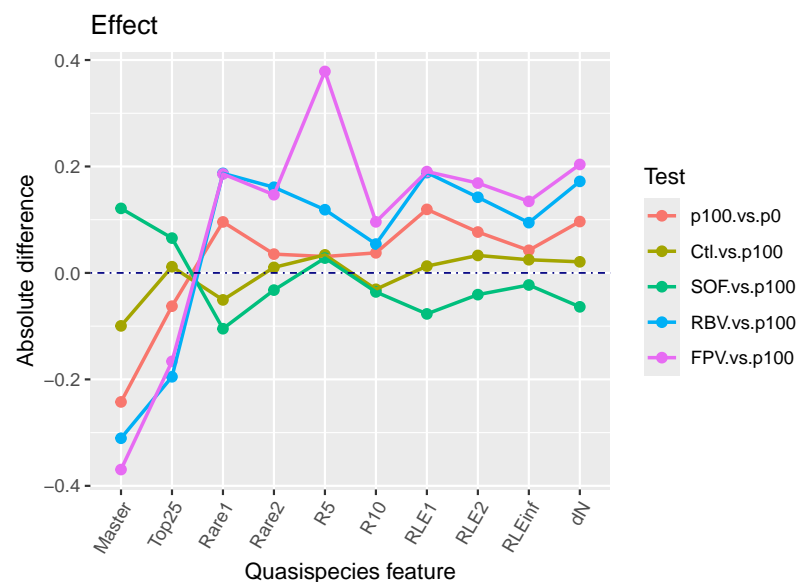
The changes in quasispecies structure indicators and in quasispecies maturity score are evaluated and tested between the following states:

- Pass 0 (p0) quasispecies versus pass 100 (p100), with no treatment in between (Supplementary Table S3).
- Pass 110 (p110, Ctl), with extra 10 passages to p100, versus p100, with p100 used as treatment baseline (Supplementary Table S4)

- Ten passages of treatment with sofosbuvir (SOF), starting at p100, versus the base line p100 (Supplementary Table S5).
- Ten passages of treatment with ribavirin (RBV), starting at p100, versus the base line p100 (Supplementary Table S6).
- Ten passages of treatment with favipiravir (FPV), starting at p100, versus the base line p100 (Supplementary Table S7).

From these tests we obtain  $p$ -values and effect sizes of different treatments over ten passages, starting at p100, on quasispecies structure indicators. Note that virtually all adjusted  $p$ -values across all comparisons are extremely low due to the very large sample sizes involved. The lowest  $t$  statistic (in absolute value) among all tests performed is 5.79 (see Supplementary Tables S3–S7), making the  $p$ -value in this context almost meaningless. Correspondingly, the 99.9% confidence intervals are very narrow, as illustrated in Supplementary Figure S4 for  $d_N$ .

Effect sizes provide better alternatives for interpreting test results. Figure 2 shows the observed differences between quasispecies pairs, whereas Figure 3 and Supplementary Figure S1 display the resulting Cohen's  $d$  values for each treatment, indicating effect magnitudes. Supplementary Figures S2 and S3 present these effects as PS and NOV values, respectively.



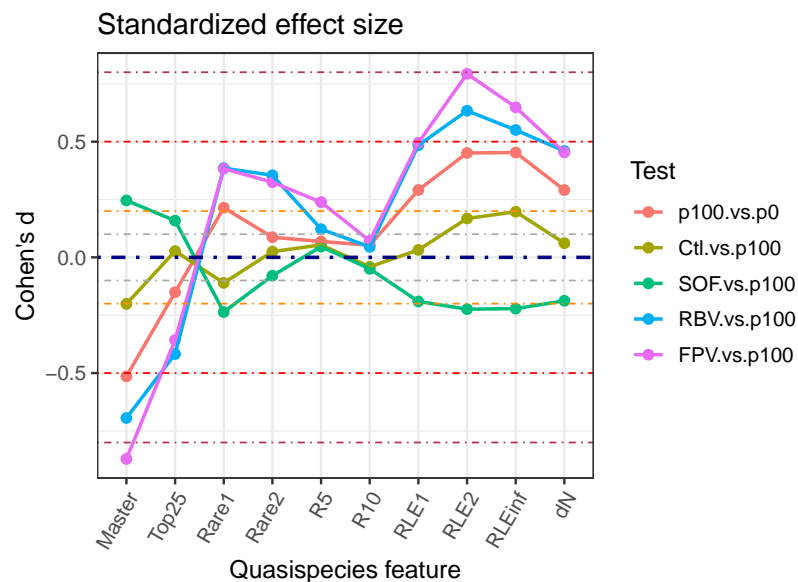
**Figure 2.** Observed differences in quasispecies indicators between biological conditions.

Let us take as reference the changes observed in 100 passages, from HCV p0 to HCV p100. All differences of indicators are highly statistically significant, with absolute  $t$  values above 12.3, and practically null  $p$ -values (Supplementary Table S3).

The effect size values for  $d_N$  gives the magnitude of the effect of treatment on all combined quasispecies indicators. Supplementary Table S8 displays the test results on  $d_N$ , where all adjusted  $p$ -values are 0. Supplementary Figure S5 depicts these values for each treatment in a barplot, with the ratio to the p100.vs.p0  $d$  value shown inside the bars, where the magnitude of each treatment effect is referred to that of the natural evolution from HCV p0 to HCV p100.

Sofosbuvir is an inhibitor, a direct acting antiviral (DAA), Ribavirin and Favipiravir act as mutagens, increasing the virus replication error rate. The mutagens acted as accelerators of quasispecies evolution without causing error catastrophe or extinction, as viral titer was maintained and even slightly increased throughout the treatments. The genetic diversification, compatible with viral functionality, achieved in ten passages with muta-

genic treatment was higher than that observed in 100 passages from HCV p0 to HCV p100, with an effect size on the maturity score  $dN$  over 1.5 times that observed in the natural evolution from HCV p0 to HCV p100 (Supplementary Figure S5). Although the Ribavirine and Favipiravir dosage was sufficient to drive HCV p0 to extinction within five passages as previously published in [22], the increased fitness achieved by HCV p100 [22,33] was enough to counteract replication inhibition. The sustained viral titer, combined with the elevated error rate, enabled the production and selection of multiple variants with comparable fitness, resulting in a more mature quasispecies.



**Figure 3.** Cohens' d standardized effect size of treatments. Dash-dot lines: effect magnitude borders for tiny 0.0–0.1, very small 0.1–0.2, small 0.2–0.5, and moderate 0.5–0.8.

The treatment with sofosbuvir resulted in a contraction of the genetic diversity, with increased top haplotypes dominance, and reduced fraction of reads for haplotypes below 1% frequency. The contraction in  $dN$  was moderate. Note that all other treatments resulted in increases of  $dN$  (Supplementary Figure S5, and Supplementary Table S8).

#### 4. Discussion

Quasispecies fitness has been shown as a determinant of lower sensitivity and failure to antiviral treatments, beyond the presence of resistance mutations, both in HCV cell culture experiments [21–24] and in HCV chronically infected patients [7]. Furthermore ribavirine has been shown to induce highly resistant quasispecies in immunocompromised patients chronically infected with HEV, with no response to increased doses [5,6]. We developed a statistical framework based on the Delta method and the asymptotic normality of multinomial parameters, and implemented it with R code, allowing to assess changes in quasispecies genetic structure indicators from single quasispecies observations sequenced at high depth. The quasispecies maturity state, understood as a surrogate of quasispecies fitness, may be taken as a general biomarker of resilience to antiviral treatments.

This study provides a robust analytical framework for assessing the structure and maturation of RNA virus quasispecies from single high-depth sequencing observations, overcoming the limitations of traditional resampling-based statistical methods in this field. By leveraging the delta method for analytical variance estimation, we enable meaningful comparisons of quasispecies diversity and structure in scenarios where replicate samples are unavailable, a common situation in virological studies. Like in -omic experiments, to mitigate the impact of sources of variability external to sampling itself—such as batch

effects, confounding factors, and technical variability in general—careful experimental design is essential [37–39]. This includes balanced blocking, randomization, and parallel processing within a short time window whenever possible. Ideally, all quasispecies to be compared should be sequenced in the same instrument run, or in balanced runs under the conditions being compared.

Our results offer clear empirical support for the anticipated effects of different evolutionary pressures on quasispecies structure. As expected, prolonged evolution in a non-coevolving environment (100 passages, HCV p0 to p100) led to increased genetic diversity, reduced dominance of the master haplotype, and greater evenness among haplotype frequencies. These changes are consistent with the theoretical understanding that high mutation rates and relaxed selective constraints facilitate the exploration of sequence space, promoting the emergence and coexistence of multiple functional variants. This genomic plasticity enhances the viral capacity for persistence and transmission, by enabling the virus to evade selective pressures, particularly immune responses, antiviral treatments, and other host-related constraints. This strategy not only complicates eradication efforts but may also promote the evolution of escape mutants, ultimately undermining therapeutic efficacy and long-term immune control. The observed increase in fitness and quasispecies maturity aligns with previous findings, emphasizing the adaptability and evolutionary potential of RNA viruses under stable conditions.

In contrast, treatment with the polymerase inhibitor sofosbuvir (SOF) produced a marked contraction in genetic diversity, as reflected by the inverse changes in all diversity and evenness indicators. This reduction in diversity, despite stable viral titers, suggests that SOF imposes strong selective pressure, favoring the survival of a subset of fit variants with low susceptibility and response to the treatment. The maintenance of viral load under these conditions highlights the capacity of highly adapted quasispecies to withstand antiviral intervention, likely through the rapid selection of pre-existing or emergent resistant haplotypes.

Mutagenic treatments with ribavirin (RBV) and favipiravir (FPV) had the opposite effect, accelerating quasispecies maturation in 10 passages beyond that observed during 100 passages of natural evolution. Both mutagens significantly increased diversity and evenness, with FPV exerting the strongest effect. Notably, these treatments did not lead to error catastrophe or viral extinction in the high-fitness HCV p100 background, underscoring the resilience of mature quasispecies populations. The ability of these populations to maintain or even increase viral titers under elevated mutational pressure suggests that high fitness confers a buffer against mutagen-induced deleterious effects, enabling the continued generation and selection of fit variants.

The goal of treating a viral infection with mutagenic agents is to achieve a cure through lethal mutagenesis [40]. During lethal mutagenesis, there is typically an initial increase in genetic diversity, followed by a sharp decline in viral titer as the viral population approaches extinction. This occurs because the mutation rate surpasses the error threshold, leading to a loss of genetic information and viability. Consequently, deleterious mutations accumulate, ultimately causing population collapse [40–42]. One of the HCV *in vitro* studies from which we have reanalyzed data reported that high-fitness hepatitis C virus, such as HCV p100, was resistant to lethal mutagenesis [22]. Furthermore, in two recent studies [5,6], we observed that repeated unsuccessful treatments with ribavirin in two chronically infected, immunosuppressed HEV patients resulted in highly mature quasispecies that were resilient to the treatment. This suggests that under mutagenic treatment, specially in monotherapy, quasispecies may evolve in one of two opposing directions [4], depending on their global fitness, genetic and phenotypic distribution structure, viral load or population size, and treatment regimen: 1—towards lethal mutagenesis and ex-

tion, or 2—towards enhanced quasispecies maturity with higher fitness and reduced susceptibility to further treatments [4,6,25].

It should be noted that the cell-culture data presented here are intended solely as an illustration of the method's applicability. Although the observed trends are in agreement with results reported in clinical datasets, these findings alone are not sufficient to draw definitive conclusions. Additional investigations, including biological replicates of cell-culture experiments, will be required to strengthen the evidence. The overall patterns appear consistent, but the magnitude of the effects may differ across independent studies.

## 5. Conclusions

From a methodological perspective, our approach addresses a critical gap in the analysis of viral quasispecies, where single-sample comparisons are often the only option. However, it is important to note that these samples, despite being single observations, represent a quasispecies with a very high sample size,  $n$ , as the number of sequenced reads. The analytical variance estimates derived via the delta method offer a reliable basis for hypothesis testing and confidence interval construction, free from the biases and limitations inherent in bootstrap and jackknife resampling for quasispecies diversity indices.

In summary, this work advances the quantitative analysis of viral quasispecies by introducing a statistically rigorous and biologically insightful framework for single-sample comparisons. Our findings underscore the dynamic interplay between mutation, selection, and treatment in shaping viral populations, with important implications for understanding viral evolution, adaptation, and the emergence of resistance to antiviral therapies. Future research should aim to validate and extend these observations, as well as investigate the framework's potential utility in clinical settings, where monitoring quasispecies diversity may help guide treatment strategies and predict therapeutic outcomes.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms13092029/s1>, **Supplementary File S1. Supplementary methods and results.** A supplementary document with methods, equations, results, tables, and figures, including the R code, required libraries and versions. Follows a list with captions of tables and figures in this Supplementary File. **Table S1. Cohen's  $d$  interpretation guidelines.** **Table S2. Variance of rarefied maturity indicators.** **Table S3. Test results.** p100 quasispecies versus p0. Dif: p100-p0, Ratio: p100/p0, adj.pval: BH adjusted  $p$ -values,  $d$ : Effect size, NOV: Non-overlap, PS: Probability of superiority. **Table S4. Test results.** p110 (Ctl) quasispecies versus p100. Dif: p110-100, Ratio: p110/p100. **Table S5. Test results.** Effect of 10 passes treated with Sofosbuvir, starting at p100. Dif: SOF-p100, Ratio: SOF/p100. **Table S6. Test results.** Effect of 10 passes treated with Ribavirine, starting at p100. Dif: RBV-p100, Ratio: RBV/p100. **Table S7. Test results.** Effect of 10 passes treated with Favipiravir, starting at p100. Dif: FPV-p100, Ratio: FPV/100. **Table S8. Summary of treatment results on  $d_N$ ,** with  $dN.d.R$  the ratio of  $d_N d$  values to the p100.vs.p0  $d_N d$  value. **Figure S1. Standardized effect size.** Cohens'  $d$  standardized effect size of treatments. Background color depicts effect magnitude regions. Gray: tiny effect, orange: very small effect, pink: small effect, purple: moderate effect, white: large effect. **Figure S2. Probability of superiority.** Standardized effect size of treatments, evaluated in probabilistic terms as probability of superiority. **Figure S3. Distributions non-overlap.** Standardized effect size of treatments, evaluated in probabilistic terms as distributions non-overlap. **Figure S4. Distance from state A.** Maturity score as the normalized distance of the quasispecies from state A. With 99.9% CIs. Dash-dot line at p100  $d_N$  as reference. **Figure S5. Effect size on  $d_N$  in each treatment.** Numbers inside bars:  $dN.d.R$ , ratio of  $d_N d$  values to p100.vs.p0  $d$  value. **Supplementary File S2. Fasta files zipped folder.** Zipped folder with the fasta files of all quasispecies in the study. Each sequence is a quasispecies haplotype, with headers expressing IDs, and frequencies as read counts and percentage. Refs. [6,30,43] are cited within.

**Author Contributions:** J.G. and M.S. conceptualization, development of the equations, and wrote the initial paper draft, J.G. and J.Q. further developed the manuscript and prepared the tables and figures. J.G. and M.I.-L. developed the software and the formal analysis. S.C.-C., C.C., and A.G.-C. worked on figures, results, and corrections of the main manuscript. J.Q. validation, supervision, and funding acquisition. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was partially supported by Grant PI22/00258 funded by Instituto de Salud Carlos III (ISCIII) and Grant PMPER24/00018 both under the Ministry of Science, Innovation and Universities of Spain and cofounded by the European Union; M.S. received support from the Spanish Ministry of Education Grant PID2023-148013OB-C22 and Generalitat de Catalunya Grant 2021 SGR 01421. C.C. received a predoctoral fellowship from the Vall d’Hebron Institute of Research (VHIR). M.I.-L. was supported by a fellowship from the “La Caixa” Foundation (ID 100010434), with code “LCF/BQ/DR23/12000020”. S.C.-C. received support from the Spanish Ministry of Education grant FPU21/04150.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Sequence data that support the findings of this study was published in [22] and are openly available in the GenBank Sequence Read Archive database with BioProject “Resistance of high fitness hepatitis C virus to lethal mutagenesis” and accession PRJNA720288. The FASTA files with haplotypes and frequencies of each sample used in the study are given in a zipped file as Supplementary Material.

**Acknowledgments:** The authors thank Julie Sheldon, Mariu Soria, Celia Perales, and Esteban Domingo for sharing the biological data used in this study, which was obtained through a previous collaborative project. Likewise, the authors are grateful to the anonymous reviewers for their valuable suggestions. We also thank Roche Diagnostics S.L.U. for supporting the APC charges and covering the full cost of publishing this manuscript.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Domingo, E.; Perales, C. Viral quasispecies. *PLoS Genet.* **2019**, *15*, e1008271. [[CrossRef](#)]
2. Domingo, E.; Sheldon, J.; Perales, C. Viral Quasispecies Evolution. *Microbiol. Mol. Biol. Rev.* **2012**, *76*, 159–216. [[CrossRef](#)]
3. Domingo, E.; García-Crespo, C.; Perales, C. Historical Perspective on the Discovery of the Quasispecies Concept. *Annu. Rev. Virol.* **2021**, *8*, 51–72. [[CrossRef](#)]
4. Todt, D.; Walter, W.; Brown, R.; Steinmann, E. Mutagenic Effects of Ribavirin on Hepatitis E Virus—Viral Extinction versus Selection of Fitness-Enhancing Mutations. *Viruses* **2016**, *8*, 283. [[CrossRef](#)]
5. Gregori, J.; Colomer-Castell, S.; Campos, C.; Ibañez-Llagoña, M.; Garcia-Cehic, D.; Rando-Segura, A.; Adombie, C.; Pintó, R.; Guix, S.; Bosch, A.; et al. Quasispecies Fitness Partition to Characterize the Molecular Status of a Viral Population. Negative Effect of Early Ribavirin Discontinuation in a Chronically Infected HEV Patient. *Int. J. Mol. Sci.* **2022**, *23*, 14654. [[CrossRef](#)]
6. Colomer-Castell, S.; Gregori, J.; Garcia-Cehic, D.; Riveiro-Barciela, M.; Buti, M.; Rando-Segura, A.; Vico-Romero, J.; Campos, C.; Ibañez-Llagoña, M.; Adombi, C.M.; et al. In-Host HEV Quasispecies Evolution Shows the Limits of Mutagenic Antiviral Treatments. *Int. J. Mol. Sci.* **2023**, *24*, 17185. [[CrossRef](#)] [[PubMed](#)]
7. Gregori, J.; Ibañez-Llagoña, M.; Colomer-Castell, S.; Garcia-Cehic, D.; Campos, C.; Quer, J. Association of liver damage and quasispecies maturity in chronic HCV patients: The fate of a quasispecies. *Microorganisms* **2024**, *12*, 2213. [[CrossRef](#)]
8. Salicru, M.; Menendez, M.; Morales, D.; Pardo, L. Asymptotic distributions of (h,Fi)-entropies. *Commun. Stat. Theory Methods* **1993**, *22*, 2015–2031. [[CrossRef](#)]
9. Pardo, I.; Morales, D.; Salicrú, M.; Menéndez, M. Large sample behavior of entropy measures when parameters are estimated. *Commun. Stat.-Theory Methods* **1997**, *26*, 483–501. [[CrossRef](#)]
10. Rey, A.; Frery, A.; Lucini, M.; Gambini, J.; Chagas, E.; Ramos, H. Asymptotic Distribution of Certain Types of Entropy under the Multinomial Law. *Entropy* **2023**, *25*, 734. [[CrossRef](#)] [[PubMed](#)]
11. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 1994. [[CrossRef](#)]
12. Hesterberg, T.; Moore, D.; Monaghan, S.; Clipson, A.; Epstein, R.; G.P., M., *Bootstrap Methods and Permutation Tests*. In *Introduction to the Practice of Statistics*, 5th ed.; McCabe, W., Ed.; Freeman and Co.: New York, NY, USA, 2005; Chapter 14, pp. 14.1–14.70.

13. Ramachandran, K.; Tsokos, C. Chapter 13—Empirical methods. In *Mathematical Statistics with Applications in R (Third Edition)*, 3rd ed.; Ramachandran, K.; Tsokos, C., Eds.; Academic Press: Cambridge, MA, USA, 2021; pp. 531–568. [[CrossRef](#)]
14. Gregori, J.; Ibañez-Lligoña, M.; Colomer-Castell, S.; Campos, C.; Quer, J. Virus Quasispecies Rarefaction: Subsampling with or without Replacement? *Viruses* **2024**, *16*, 710. [[CrossRef](#)]
15. Shao, J.; Wu, C.J. A general theory for jackknife variance estimation. *Ann. Stat.* **1989**, *17*, 1176–1197. [[CrossRef](#)]
16. Shao, J. *Consistency of Jackknife Variance Estimators*; Technical Report 88-58; Purdue University, Department of Statistics: West Lafayette, IN, USA, 1988.
17. Shao, J. Consistency of jackknife variance estimators jun. *Statistics* **1991**, *22*, 49–57. [[CrossRef](#)]
18. Bera, A.; Koley, M. A History of the Delta Method and Some New Results. *Sankhya B* **2023**, *85*, 272–306. [[CrossRef](#)]
19. Hosmer, D.; Lemeshow, S.; May, S., Appendix 1: The Delta Method. In *Applied Survival Analysis: Regression Modeling of Time-to-Event Data*; John Wiley and Sons, Inc.: Hoboken, NJ, USA, 2008; pp. 355–358.
20. VerHoef, J. Who invented the Delta Method? *Am. Stat.* **2012**, *66*, 124–127. [[CrossRef](#)]
21. Gallego, I.; Sheldon, J.; Moreno, E.; Gregori, J.; Quer, J.; Esteban, J.; Rice, C.; Domingo, E.; Perales, C. Barrier-Independent, Fitness-Associated Differences in Sofosbuvir Efficacy against Hepatitis C Virus. *Antimicrob Agents Chemother* **2016**, *60*, 3786–3793. [[CrossRef](#)] [[PubMed](#)]
22. Gallego, I.; Gregori, J.; Soria, M.; García-Crespo, C.; García-Álvarez, M.; Gómez-González, A.; Valiergue, R.; Gómez, J.; Esteban, J.; Quer, J.; et al. Resistance of high fitness hepatitis C virus to lethal mutagenesis. *Virology* **2018**, *523*, 100–109. [[CrossRef](#)]
23. Moreno, E.; Gallego, I.; Gregori, J.; Lucía-Sanz, A.; Soria, M.; Castro, V.; Beach, N.; Manrubia, S.; Quer, J.; Esteban, J.; et al. Internal Disequilibria and Phenotypic Diversification during Replication of Hepatitis C Virus in a Noncoevolving Cellular Environment. *J. Virol.* **2017**, *91*, e02505-16. [[CrossRef](#)] [[PubMed](#)]
24. García-Crespo, C.; Gallego, I.; Soria, M.; de Ávila, A.; Martínez-González, B.; Vázquez-Sirvent, L.; Lobo-Vega, R.; Moreno, E.; Gómez, J.; Briones, C.; et al. Population Disequilibrium as Promoter of Adaptive Explorations in Hepatitis C Virus. *Viruses* **2021**, *13*, 616. [[CrossRef](#)] [[PubMed](#)]
25. Gregori, J.; Colomer-Castell, S.; Ibañez-Lligoña, M.; Garcia-Cehic, D.; Campos, C.; Buti, M.; Riveiro-Barciela, M.; Andrés, C.; Piñana, M.; González-Sánchez, A.; et al. In-Host Flat-like Quasispecies: Characterization Methods and Clinical Implications. *Microorganisms* **2024**, *12*, 1011. [[CrossRef](#)]
26. Duris, F.; Gazdarica, J.; Gazdaricova, I.; Strieskova, L.; Budis, J.; Turna, J.; Szemes, T. Mean and variance of ratios of proportions from categories of a multinomial distribution. *J. Stat. Distrib. Appl.* **2018**, *5*, 2. [[CrossRef](#)]
27. Kim, H. Statistical notes for clinical researchers: Effect size. *Restor Dent. Endod.* **2015**, *40*, 328–331. [[CrossRef](#)]
28. Nakagawa, S.; Cuthill, I. Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biol. Rev. Camb. Philos. Soc.* **2007**, *82*, 591–605. . Erratum in: *Biol. Rev. Camb. Philos. Soc.* **2009**, *84*, 515. [[CrossRef](#)]
29. Kelley, K.; Preacher, K. On effect size. *Psychol. Methods* **2012**, *17*, 137–152. [[CrossRef](#)] [[PubMed](#)]
30. Sawilowsky, S. New Effect Size Rules of Thumb. *J. Mod. Appl. Stat. Methods* **2009**, *8*, 26. [[CrossRef](#)]
31. Marukian, S.; Jones, C.; Andrus, L.; Evans, M.; Ritola, K.; Charles, E.D.; Rice, C.M.; Dustin, L.B. Cell culture-produced hepatitis C virus does not infect peripheral blood mononuclear cells. *Hepatology* **2008**, *48*, 1843–1850. [[CrossRef](#)]
32. Perales, C.; Beach, N.; Gallego, I.; Soria, M.; Quer, J.; Esteban, J.; Rice, C.; Domingo, E.; Sheldon, J. Response of Hepatitis C Virus to Long-Term Passage in the Presence of Alpha Interferon: Multiple Mutations and a Common Phenotype. *J. Virol.* **2013**, *87*, 7593–7607. [[CrossRef](#)]
33. Sheldon, J.; Beach, N.; Moreno, E.; Gallego, I.; Piñeiro, D.; Martínez-Salas, E.; Gregori, J.; Quer, J.; Esteban, J.; Rice, C.; et al. Increased replicative fitness can lead to decreased drug sensitivity of hepatitis C virus. *J. Virol.* **2014**, *88*, 12098–12111. [[CrossRef](#)]
34. Gallego, I.; Soria, M.; Gregori, J.; Quer, J.; Esteban, J.; Rice, C.; Domingo, E.; Perales, C. Lethal Mutagenesis of Hepatitis C Virus Induced by Favipiravir. *PLoS ONE* **2016**, *11*, e0164691. [[CrossRef](#)]
35. Ortega-Prieto, A.; Sheldon, J.; Grande-Pérez, A.; Tejero, H.; Gregori, J.; Quer, J.; Esteban, J.; Domingo, E.; Perales, C. Extinction of Hepatitis C Virus by Ribavirin in Hepatoma Cells Involves Lethal Mutagenesis. *PLoS ONE* **2013**, *8*, e71039. [[CrossRef](#)]
36. Gregori, J.; Soria, M.; Gallego, I.; Guerrero-Murillo, M.; Esteban, J.; Quer, J.; Perales, C.; Domingo, E. Rare haplotype load as marker for lethal mutagenesis. *PLoS ONE* **2018**, *13*, e0204877. [[CrossRef](#)]
37. Montero, O.; Hedeland, M.; Balgoma, D. Trials and tribulations of statistical significance in biochemistry and omics. *Trends Biochem. Sci.* **2023**, *48*, 503–512. [[CrossRef](#)]
38. Forshed, J. Experimental Design in Clinical 'Omics Biomarker Discovery. *J. Proteome. Res.* **2017**, *16*, 3954–3960. [[CrossRef](#)] [[PubMed](#)]
39. Singh, P.; Benayoun, B. Considerations for reproducible omics in aging research. *Nat. Aging* **2023**, *3*, 921–930. [[CrossRef](#)] [[PubMed](#)]
40. Domingo, E. Viruses at the edge of adaptation. *Virology* **2000**, *270*, 251–253. [[CrossRef](#)]
41. Perales, C.; Gallego, I.; de Ávila, A.; , Soria, M.; Gregori, J.; Quer, J.; Domingo, E. The increasing impact of lethal mutagenesis of viruses. *Future Med. Chem.* **2019**, *11*, 1645–1657. [[CrossRef](#)]

42. Hassine, I.; M'hadheb, M.; Menéndez-Arias, L. Lethal Mutagenesis of RNA Viruses and Approved Drugs with Antiviral Mutagenic Activity. *Viruses* **2022**, *14*, 841. [[CrossRef](#)]
43. Magoc, T.; Salzberg, S.L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **2011**, *27*, 2957–2963. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.