Video Article

# Isolation of Fidelity Variants of RNA Viruses and Characterization of Virus Mutation Frequency

Stéphanie Beaucourt[1], Antonio V. Bordería[1], Lark L. Coffey[1], Nina F. Gnädig[1], Marta Sanz-Ramos[1], Yasnee Beeharry[1], Marco Vignuzzi[1]

[1]Viral Populations and Pathogenesis lab and CNRS 3015, Institut Pasteur

Correspondence to: Marco Vignuzzi at marco.vignuzzi@pasteur.fr

URL: https://www.jove.com/video/2953
DOI: doi:10.3791/2953

Keywords: Immunology, Issue 52, Polymerase fidelity, RNA virus, mutation frequency, mutagen, RNA polymerase, viral evolution

Date Published: 6/16/2011

Citation: Beaucourt, S., Bordería, A.V., Coffey, L.L., Gnädig, N.F., Sanz-Ramos, M., Beeharry, Y., Vignuzzi, M. Isolation of Fidelity Variants of RNA Viruses and Characterization of Virus Mutation Frequency. *J. Vis. Exp.* (52), e2953, doi:10.3791/2953 (2011).

## Abstract

RNA viruses use RNA dependent RNA polymerases to replicate their genomes. The intrinsically high error rate of these enzymes is a large contributor to the generation of extreme population diversity that facilitates virus adaptation and evolution. Increasing evidence shows that the intrinsic error rates, and the resulting mutation frequencies, of RNA viruses can be modulated by subtle amino acid changes to the viral polymerase. Although biochemical assays exist for some viral RNA polymerases that permit quantitative measure of incorporation fidelity, here we describe a simple method of measuring mutation frequencies of RNA viruses that has proven to be as accurate as biochemical approaches in identifying fidelity altering mutations. The approach uses conventional virological and sequencing techniques that can be performed in most biology laboratories. Based on our experience with a number of different viruses, we have identified the key steps that must be optimized to increase the likelihood of isolating fidelity variants and generating data of statistical significance. The isolation and characterization of fidelity altering mutations can provide new insights into polymerase structure and function[1-3]. Furthermore, these fidelity variants can be useful tools in characterizing mechanisms of virus adaptation and evolution[4-7].

## Video Link

The video component of this article can be found at https://www.jove.com/video/2953/

## Protocol

### 1. Determine the range of mutagen concentrations that is minimally toxic to cells

The purpose of this exercise is to determine what range of mutagen concentrations can be used during an infection without excess cell toxicity. Essentially, you will want to reproduce the conditions that will be required for virus infection. For most viruses, infections last between 2 and 7 days. Prepare enough plates to sample cells at each day. If non-adherent cells are used, modify the protocol accordingly.

1. On the day before the experiment, seed 7 x $10^5$ HeLa cells/well in a 6-well plate that achieve a sub-confluent (75%) monolayer the day of the experiment. Each well of the plate will be treated with a different concentration of mutagen, permitting a range of 6 concentrations.
2. On the day of the experiment, prepare mutagen dilutions in tissue culture medium. For HeLa cells, use a range of 0 to 1000 µM for base analogs (ribavirin, 5-fluorouracil, 5-azacytidine), 0 to 50 mM for $MgCl_2$, and 0 to 5 mM for $MnCl_2$.
3. Aspirate the medium from the wells and replace with 2 ml mutagen-supplemented medium and return to incubator.
4. Every 24 hours, use one plate of cells to check cell viability. This can be accomplished by performing trypan blue staining or using commercialized fluorescence/luminescence assays (e.g. Promega's CellTiter-Glo® Luminescent Cell Viability Assay).
5. For trypan blue exclusion staining, detach cells from one plate (treated with different concentrations of mutagen) and gently pellet cells by centrifugation.
6. Discard supernatant and resuspend cells in PBS (serum can interfere with staining).
7. Mix 1 volume cell suspension in PBS with 1 volume of 0.4% trypan blue, incubate for 2 minutes at room temperature.
8. In a hemacytometer, count the viable (unstained) and non-viable (blue stained) cells. Calculate the percentage of viable cells for each concentration of mutagen, including the untreated control. We find that conditions that result in less than 50% cell death by the infection endpoint (when maximum titers are reached) are ideal for isolation of mutagen resistance.

### 2. Determine the optimal non toxic mutagen concentration that moderately reduces virus titers (approximately 0.5-2 log reduction)

This exercise serves to determine the concentration of mutagen that will exert a strong selective pressure, without over-mutagenizing the population. For RNA mutagens, we find that this corresponds to 10.5-2 logs reductions in virus titer. At these concentrations, every genome is mutated in at least one or two positions. The mutagen may aid in the generation of the resistance mutation, which will then be selected

over passage. If too many mutations are introduced (at very high mutagen concentrations), the mutants bearing the resistance mutation will themselves be lethally mutagenized, impeding their isolation.

1. Seed plates with cells using the same conditions as for step 1
2. On the day of the experiment, prepare mutagen dilutions in tissue culture medium. Use the same range of concentrations as determined above, but exclude concentrations that resulted in over 50% cell death. Prepare enough medium to cover each well twice (4 ml per well), allowing for a pretreatment of cells before infection, with each concentration of mutagen.
3. Aspirate the medium and pretreat cells with mutagen by incubating in mutagen-supplemented medium for 2 hours. For most cell types, this is ample time for uptake of mutagen.
4. Remove medium and infect with virus at low multiplicity of infection MOI (0.1 or 0.01) in a minimum volume (200 µl for 6-well plates). Incubate 15-60 minutes to allow virus to infect cells. Rock the plate at regular intervals to ensure that the inoculum covers the cell monolayer.
5. Aspirate the inoculated virus and wash twice with 2 ml PBS to remove as much of the inoculum as possible.
6. Add medium with the appropriate concentrations of mutagen to each well and incubate cells for the equivalent of 3-6 replication cycles.
7. Harvest virus from each well and determine the antiviral effect on virus titer. This can be done by standard plaque assay or limiting dilution ($TCID_{50}$). Note: Quantification of virus by techniques that measure only RNA synthesis may not be suitable because the mutagenic effects may not be detected. Mutagenized genomes containing lethal mutations may still be detected by qRT-PCR for example, but would not be observed in virus viability assays.
8. From the calculated titers, identify the concentration of mutagen that reduces virus titers (compared to the untreated control infection) by 10.5-2 logs that is also not highly toxic to cells (ideally, less than 50% toxicity).

## 3. Isolation and identification of mutagen resistant variants

Perform large population size passages in the optimal mutagen concentration defined above and check virus titers across the passage series. As a control, passage virus in growth medium without any mutagen. As another control to monitor the potential emergence of defective interfering particles (DI), perform fresh infections in absence of mutagen at each passage step (unpassaged control).

1. On the day before infection, seed 25 $cm^2$ flasks with 1.5 x $10^6$ HeLa cells (other flask sizes can be used) to obtain sub-confluent monolayers the following day.
2. On the day of the infection, pretreat cells for 2 hours with medium containing the optimal concentration of each mutagen, determined in section 2.
3. Remove medium and infect cells with a minimum volume at a MOI of 1 or the largest MOI that does not result in defective interfering (DI) particle formation for the virus being studied.
4. After 30-60 minutes of infection, aspirate inoculum and wash twice with PBS, then add fresh medium supplemented with mutagen at the appropriate concentrations.
5. Incubate for the period of time determined in sections 1 and 2 that corresponds to maximum virus titer in these conditions. Harvest the progeny virus.
6. Titer virus at each passage and repeat the 3 preceding steps.
7. During the first few passages, virus titers of mutagen treated samples should drop accordingly, compared to the original virus titer and the control (untreated and unpassaged) virus titers. If virus titers of mutagen passaged samples climb back to the same levels as the untreated control, the population likely contains a mutagen resistant variant. Up to 20 or 30 passages may be required, although we isolated most of our fidelity variants between 5 and 15 passages.
8. Once virus titers for a given passage series reach the same magnitude as the untreated control titers, extract RNA from all samples including the untreated control from the same passage number. RNA extraction kits or Trizol extraction can be used.
9. Perform RT-PCR using primers that amplify the polymerase or replicase genes of the virus of interest. In a second step, the entire genome (at least coding regions) should be sequenced to examine whether resistance phenotypes map to other virus genes as well. This is particularly important for base analog mutagens, such as ribavirin, which affects other aspects of virus and cell function. In this case the variant may be resistant to one of these other antiviral activities and will not be a fidelity variant.
10. Purify the RT-PCR products using a PCR purification kit and sequence to obtain the mutagen-resistant population consensus sequence. Include background controls for sequencing error (see Discussion).
11. Using a sequence alignment software and the virus consensus sequence as a reference, align the sequences. Identify the new point mutations, with particular attention to any that appear exclusively in the mutagen treated population at the passage where virus titers reach normal levels. If this mutation is not present in earlier passages and not present in untreated controls from the same passage number (indicating adaptation to cell culture passage), then this mutation is likely responsible, at least in part, for mutagen resistance. Do not rely on base-called sequence (the text version of sequence) and alignment software alone. Check the chromatograms for minority peaks that may have been missed by alignment software. A mutant representing 20-30% of the total population will still show as a peak, but too small to be identified as an 'N' by standard sequence analysis.

## 4. Once a mutation has been identified, isolate or generate the variant and confirm the resistance phenotype to several RNA mutagens

Next, the variant presenting the identified mutation is isolated to confirm its link to the resistance phenotype. It is essential that the mutation suspected of changing fidelity is studied in a genetically clean background (that is, not presenting additional mutations elsewhere in the genome). In the best situation, an infectious cDNA clone exists that would permit the generation of a stock of the mutagen-resistant variant by site directed mutagenesis on a clean genetic background. In this case, section 4 is not necessary. However, if a cDNA clone is not available, isolation can be done by plaque purification of virus, described below. More than one round of plaque purification may be required to isolate the variant on a clean, genetic background.

1. Isolation of the mutagen resistant mutant by plaque assay.

1. To isolate the identified mutant, perform a standard plaque assay under agarose (0.5 to 1% final wt/vol) overlay in 6-well plates. Prepare serial dilutions of the virus, based on the stock titers, to dilutions that will produce between 10 and 50 well-separated plaques.
2. When plaques are clearly visible (usually 2 to 5 days post infection, depending on the virus), mark the location of the plaques on the plates and using a p200 pipette with filter tip, gently plunge the tip through the agarose overlay being careful not to dislodge and shift the overlay's position (which would result in cross contamination of individual plaques).
3. Carefully lift the tip out of the overlay and transfer the agarose plug contained in the tip to an eppendorf containing 250 µl of medium and vortex. Do not worry if the tip that is removed does not contain agarose, for many RNA viruses, an average plaque for many RNA viruses contains $10^5$ viruses and a sufficient amount will be transferred simply by touching the tip to the surface of the plaque.
4. Pick up to 10 plaques per mutagen treatment. Depending on the chromatograms of sequencing that identified the mutation, estimate approximately what percentage of the population contains the desired mutation. The objective is to isolate three or four plaques with the mutation. Some of these mutants will also carry additional, unwanted mutations, which will later be identified by sequencing.
5. Extract the RNA from these samples (but save half of the sample to make a larger stock of virus), and perform RT-PCRs that will permit the sequencing of the entire genome. On average, RNA viruses will contain up to two mutational differences with respect to the consensus sequence; thus, sequence 3 or 4 plaque purified viruses at a time to identify the isolate that contains the desired mutation without any additional mutations.
6. Once identified, make a larger stock of this virus for all downstream studies, using the plaque purified sample obtained above to infect a larger flask of cells, e.g. $8x10^6$ HeLa cells in a T75 flask..

2. Confirm the mutagen sensitivity/resistance conferred by the identified mutation.
    1. Using the isolated or newly generated clone, and a wild type control virus prepared under similar conditions, repeat the experiments in section 2 using either a full range of mutagen concentrations, or the concentration at which the resistance mutation was generated.
    2. Use several different RNA mutagenic conditions (ribavirin, 5-fluorouracil, 5-azacytidine, increased $Mg^{2+}$, $Mn^{2+}$). If the polymerase variant is resistant to more than one kind of mutagen, then it is more likely that this variant is high fidelity. Alternatively, it is possible that a resistance mutation is specific for a single mutagenic condition, particularly since some of these compounds affect RNA viruses through a number of mechanisms [8].

## 5. Check replication rates

Since fidelity- altering mutations most often map to the polymerase, it is possible that the same polymerase mutation will significantly alter replication kinetics and it is important to determine similarities and differences in replication that will permit a better comparison of differences in mutation frequencies performed below. To do so, examine replication by at least two complimentary approaches – one that examines virus production and another that examines RNA synthesis.

1. One-step virus growth kinetics
    1. On the day before the experiment, seed 6-well plates as needed, one plate per time point to be tested. Consider using triplicate wells for each mutant and the wild type virus.
    2. On the day of the experiment, remove the medium and infect wells with each virus at a MOI of 10 to ensure that every cell is simultaneously infected. Incubate 30-60 minutes at 37°C.
    3. Rock plates every 10 minutes to avoid drying of the cell monolayer. Remove virus and wash twice with 2 ml PBS. It is important to remove as much of the inoculum as possible. Replace with growth medium.
    4. Following infection at time=0, harvest the virus from one plate. Return plates to incubator and harvest the viruses at regular intervals that span a single replication cycle (e.g. for 3h, 5h, 7h, 9h, 12h, 24h).
    5. Titer virus, harvested at each time point (e.g. plaque assay, TCID50, FFU assay), and graph the growth curves of titer versus time.

2. Kinetics of RNA synthesis
   The kinetics of RNA synthesis can be monitored using one of the approaches indicated below. If possible, the same samples used to determine one-step growth kinetics should be used to measure RNA levels.
    1. qRT-PCR. This procedure gives highly quantitative measures of replication over a large range, from a few genome copies to $>10^{10}$, depending on the sensitivity of the assay. Design primers and probes that cover a small fragment (<200 bp) of a highly conserved genomic region.
    2. Northern blot analysis. Although less quantitative than qRT-PCR, this technique permits visual confirmation that replication results in full-length genomes and that no significant chain termination occurs as a result of the polymerase mutation.
    3. Expression of a reporter gene. If a cDNA clone expressing a reporter gene (e.g. luciferase) is available, then this may be used as a surrogate to examine replicative capacity. However, the recombinant virus should not be used for other applications (such as determination of mutation frequency) since the selective pressures acting on this virus will not be the same, in particular since these viruses have a tendency to delete the inserted reporter gene.

## 6. Measure mutation frequencies

This is a critical step in confirming that the identified polymerase mutation conferring resistance to mutagen alters replication fidelity. It is important to note that the mutation *frequencies* measured here are not mutation *rates*. To determine rates, a very careful measurement of replication kinetics (amount of RNA synthesized and length of replication cycle) must be factored in. Measuring mutation frequencies however, as long as passage history and replication kinetics are monitored, provides reproducible, quantitative measures of replication fidelity. Mutation frequencies can be determined either in the viable virus population (plaque clones or limiting dilution) or in the total virus population (virus stock or supernatant). To determine mutation frequencies, prepare virus stocks from a later passage (e.g. passage 2 or beyond). It is important that the virus population has had time to expand its genetic diversity closer to a mutation-selection equilibrium.

1. Mutation frequencies of viable virus population
   This approach, although more laborious, gives information on how many mutations are present on the average genome that retains replication competence. It should be noted, however, that a bias for higher fitness variants will occur and lower fitness, viable variants that do not easily plaque, for example, may not be detected. As such, it also permits better measures of synonymous (dS) and non- synonymous (dN) nucleotide substitutions that can be used to explore whether positive selection is acting on the population. However, because fewer mutations will be quantified, a larger number of sequences will be needed for statistical analysis. This technique relies on isolation of individual viruses by either plaque purification or limiting dilution, as described above. As a starting point, we recommend isolation of 48 individual "clones" of wild type virus and the mutagen-resistant variant. Each clonal population isolated in this manner is expected to carry whatever mutation the founder genome presented. The amount of RNA present in the isolated plaque or limiting dilution well is generally sufficient for amplification by RT-PCR. If necessary, a short amplification (less than one replication cycle) on a minimum number of cells (e.g. 24 well plate format) can be used to obtain more RNA; however, minimum amplification should be performed to avoid the accumulation of new mutations. Note that each clone and population to be compared must undergo the same number of replication cycles.
   1. Isolate 24 to 48 virus clones by plaque purification or limiting dilution.
   2. Extract RNA from the isolated clonal populations
   3. RT-PCR amplify a fragment covering up to 3kb for each sample. It is best to cover the structural protein region, which tends to tolerate more viable mutations than more conserved regions of non-structural genes.
   4. Purify the PCR products, sequence and perform mutation analysis (section 7).

2. Mutation frequencies of total virus population
   Although an advantage of this second method is that even low fitness variants will be included in the sequencing, permitting a broader picture of the spectrum of mutations. However, it may not be ideal for phylogenetic analyses that assume viable virus populations (e.g. dN/dS values) and identification of the most relevant mutations, since lethal changes (altered RNA structure, stop codons, dramatic amino acid changes) cannot be entirely identified and would be retained in the analysis. Nevertheless, it this technique permits the researcher to obtain the most statistically significant data to confirm alteration of fidelity when there is a lack of an in vitro biochemical assay. This technique relies on RT-PCR amplification of total virion RNA, including genomes with low-fitness or lethal mutations that will not produce plaques. The mutation frequencies obtained by this method can be 10-fold higher than by plaque or limiting dilution cloning.
   1. Extract RNA from the total virion population
   2. RT-PCR amplify a 800 to 1200 nucleotide region from a part of the genomic coding sequence that is known to tolerate mutations and have genetic variance (e.g. structural proteins). Larger fragments will not insert readily into cloning vectors such as TopoTA and will produce insufficient numbers of transformants. Although smaller fragments are even better, the coverage of genome sequenced may be too little to obtain statistical significance. A fragment of at least 800 bp permits sequence coverage with two primers and is a good compromise between maximizing sequence coverage and minimizing sequencing costs. We find that between 70 and 100 sequences covering a 800 nucleotide region reproducibly confirms the altered fidelity of the variants we have studied in the lab. Note that other vector/cloning methods can be used with equal efficiency.
   3. Purify the RT-PCR product using a commercial kit or by standard DNA extraction/precipitation.
   4. If the RT-PCR enzymes you use do not produce A overhangs, perform a 10 minute extension by adding 1 µM ATP and Taq polymerase
   5. TopoTA clone following the manufacturer's instructions
   6. For each virus population to be studied, select 96 colonies, identified as having a positive insert by blue/white screening on XGal coated plates. Initially, test the presence of inserts for each different genome region being cloned, to confirm the validity of blue/white screening by plasmid size screening on agarose gels or single-colony PCR. Using the fragment size and conditions above, we achieve 90% positives
   7. Grow each colony in liquid broth overnight in 1 ml of LB medium in 96 well bacterial culture plates
   8. The next day prepare minipreps in 96-well format.
   9. Sequence each plate with enough primers (the primers used for RT-PCR for example, or TopoTA m13 primers) to obtain maximum coverage of the cloned segment. Perform mutation analysis (section 7).

## 7. Sequence analysis

Perform sequence analyses using a reference or consensus sequence for each population and appropriate alignment software. We recommend Lasergene or Sequencher that can readily identify SNPs with respect to consensus.

1. Align the sequences using appropriate software (e.g. Lasergene or Sequencher).
2. Discard poor quality sequences (bad base calling, too many 'N's or too short in length). Identify the nucleotide range which is covered by all sequences. Since different regions of the genome will more or less tolerate mutations, for comparison reasons it is essential that the same region is fully covered for each sequenced clone retained for analysis. Thus, if several sequence reactions are performed per clone, and one sequence fails for a given clone, discard the clone (all sequences) from the analysis.
3. Identify and count the SNPs that are different from the reference strain.
4. Calculate the mutation frequency by dividing the total number of SNPs identified by the total number of nucleotides sequenced (number of clones x length of region sequenced). Presenting this number as average number of mutations per 10K nt sequenced renders the number more user-friendly then leaving it as per nucleotide. For example, for the wild type population in Table 1, 55 mutations/121,978 total nucleotides X 10,000 = 4.51 mutations per 10,000 nucleotides sequenced.
5. If the same SNP appears on a large number of clones, present two values that include or exclude these repeated mutations. Generally, for a virus population produced from a homogenous parent (by plaque purification or from an infectious clone) and passaged only a few times in cell culture, positive selection has not exerted its effects enough to cause the accumulation of the same SNP and the mutation frequency values measured this way better reflect the error frequency of the polymerase with minimal effects of positive or purifying selection.
6. Determine the mutation distribution. Make a ranked list of the number of clones in each population that present 0, 1, 2, 3, etc... mutations in the region sequenced.

7. Calculate viral population diversity by pairwise distance comparison. From the cleaned and manually edited sequences prepare an alignment including the reference region. There are several alignment programs available: ClustalW/X (http://www.clustal.org/), MUSCLE (http://www.drive5.com/muscle/), EbioX for Mac users (http://www.ebioinformatics.org/ebiox/), etc. Make sure all your sequences are the same length, crop if necessary. It is recommended that the start of the sequences is a coding codon to facilitate posterior analysis. We suggest keeping all the alignments in fasta format, which is easily read by most software available.
8. To perform the pairwise distance analysis, calculate all possible pairwise comparisons among the sequences belonging to the same population. Average mutations found in the comparisons can then be calculated to indicate population heterogeneity.
9. Values for synonymous (dS) and non-synonymous (dN) within the population can be easily obtained. We find the MEGA software (http://www.megasoftware.net/) useful and user friendly for this kind of analysis.
10. To obtain selection direction we give two possibilities although they are not the only ones possible. 1) Obtain the ratio between dN and dS. Values higher than 1 mean positive selection. Values lower than 1 mean purifying selection. 2) Use the Datamonkey webserver (http://www.datamonkey.org/). Upload your alignments in fasta format and analyze them with the SLAC module. This will give you an estimate of dN/dS.
11. Perform statistical analyses. Depending on the quantity of data and number of sequences, a variety of tests can be used. Some studies have relied on chi square tests of total number of mutations versus total number of consensus nucleotides where all clones are combined [9]. Other studies have performed Fisher's exact tests, calculated on the number of sequences presenting mutations versus the number of sequences without mutations [10]. If enough mutational data is generated, we recommend performing a ranked sum test such as Mann Whitney U. To do so, rank the number of clones in each virus population by number of mutations present on each RNA sequenced. The Mann Whitney will then test for differences in the mutation distribution of populations. For this reason, we recommend sequencing at least 800 base pairs, to increase the likelihood of finding clones with multiple mutations. This test is robust, but requires larger sample sizes. On the other hand, it does not require the same sample size for the two populations being compared (e.g., the samples from Table 1 are $n_1$=148 and $n_2$=84).

## 8. Representative results:

The dose dependent effect of mutagen concentration on cell viability and virus viability is shown in **Figure 1.** In this example, we found that virus passage in 100 μM AZC was reduced in virus titer by the targeted 10.5-2 log, but HeLa cell viability was not negatively impacted during the 2 days required for virus infection. This pilot experiment led to the choice of 100 μM AZC concentration for serial passage of virus, to select for mutagen resistance. **Figure 2** illustrates the initial reduction in titer, followed by emergence of a mutagen resistance phenotype. During the first few passages in mutagen, as lethal mutations accumulate, a significant drop in virus titers occurs. Gradually, a mutagen-resistant variant emerges an its emergence coincides with a return to virus titers no different from untreated controls. At this stage, a large percentage of the virus population presents the resistance mutation. Sequencing of this virus population reveals the amino acid change(s) responsible. Once identified, and isolated or newly generated, the mutagen-resistant virus may be less sensitive than wild type to different RNA mutagens (base analogs of different structure, for example). **Figure 3** shows a RNA mutagen resistant Coxsackie virus B3 that titers higher than wild type in the presence of ribavirin, 5-fluorouracil, and 5-azacytidine, and high $MgCl_2$ and $MnCl_2$. Broad resistance to RNA mutagens is a strong indicator of increased replication fidelity. Verifying that replication kinetics of the fidelity variant is similar to wild type virus will aid in comparison of mutation frequencies. **Figure 4** depicts the one step-growth kinetics of a high fidelity variant compared to wild type. If replication rates and final titers are not similar, then steps should be taken to compare virus populations of similar size, that have undergone the same number of rounds of replication. The link between RNA synthesis rate and replication fidelity is not well characterized, particularly *in vivo*. A slower replication rate may result in a decreased mutation frequency (higher fidelity), although this is not an absolute rule, as shown in **Figure 4**. With the above parameters established, the mutation frequencies of the fidelity variant and wild type populations can be compared to obtain genetic confirmation of altered replication fidelity. **Figure 5** shows a sequence alignment of a wild type and high fidelity variant, with point mutations identified. The mutations are counted, ranked according to number of mutations per clone (Table 1), and represented as an average mutation frequency per population, per 10,000 nucleotides sequenced, **Figure 6**.

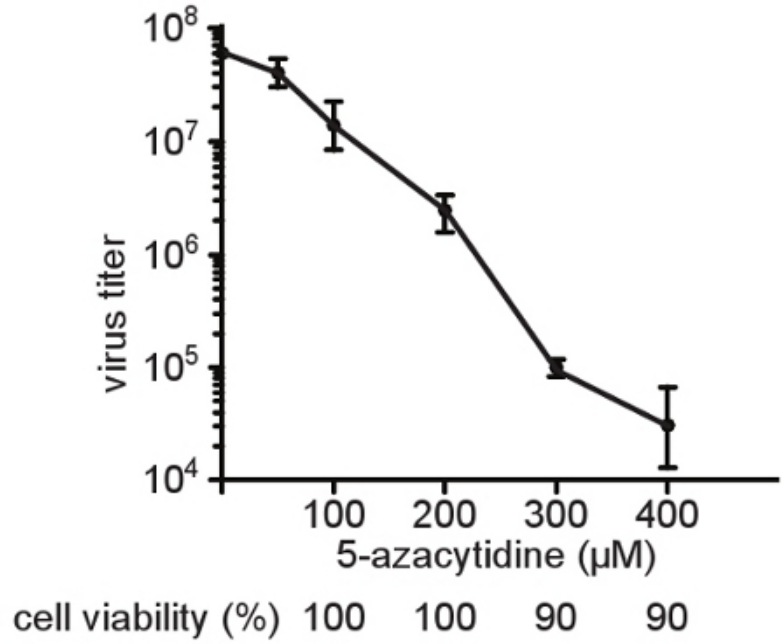


**Figure 1. Determining the optimal conditions to select for RNA mutagen resistance: retention of high cell viability with a moderate (1-2 log drop in virus titer).** HeLa cells were treated with indicated concentrations of ribavirin and infected with wild type Coxsackie virus B3 at a MOI of 0.01. 48 hours post infection, the progeny virus was harvested and titers were determined by $TCID_{50}$. The percentage of cells surviving treatment at 48 hours, determined by Trypan blue staining, is indicated below the x-axis. The results show that concentrations of 100 and 200 μM reduce virus titers by 1-2 log, without affecting cell viability.

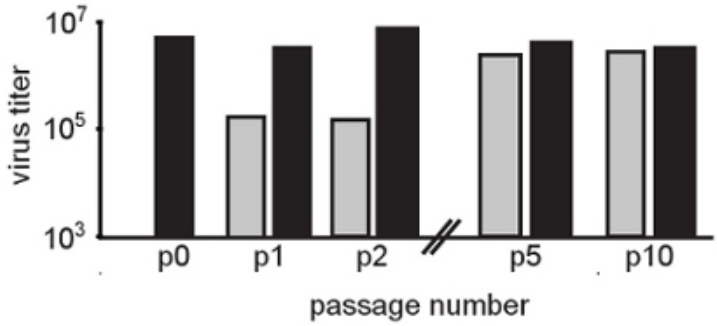


**Figure 2. Serial passage in the presence of moderate concentrations of RNA mutagens selects for mutagen resistant populations.** In this figure, Chikungunya virus was passaged in HeLa cells in the presence of 50 μM ribavirin (grey bars). Control passages were performed in absence of ribavirin (black bars). After each passage, virus progeny was quantified by classic plaque assay on BHK cells. The mutagenic effect is evident during the first passages (p1 and p2 compared to p0 starting population) where the treated virus titers drop by 2 log. Gradually, titers return to normal (untreated) levels. No significant differences are observed in passage 5 mutagen treated populations compared to untreated, suggesting that resistant variants have been selected. Indeed, consensus sequencing of the population identified unique mutations in the virus population undergoing ribavirin treatment.

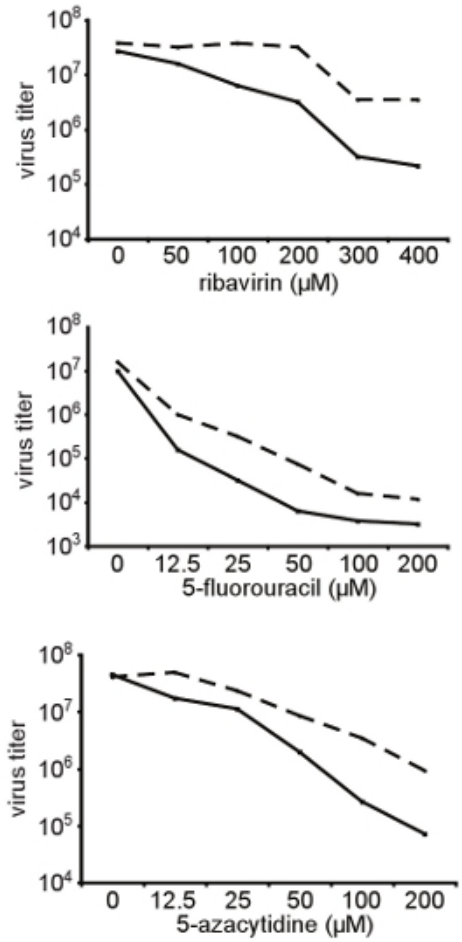


**Figure 3. Confirmation of broad resistance to RNA mutagens of different structure.** Shown here, the high fidelity A372V variant of Coxsackie virus B3 that was initially isolated in the screen described in Section 3 was generated from an infectious clone and tested for its relative sensitivity to different concentrations of different RNA mutagens (ribavirin, 5-fluorouracil, 5-azacytidine). HeLa cells were treated with indicated concentrations of ribavirin and infected with wild type Coxsackie virus B3 at a MOI of 0.01. 48 hours post infection, the progeny virus was harvested and titers were determined by $TCID_{50}$. Shown here are the titers of wild type (solid lines) and A372V variant (dashed lines) as a function of mutagen concentration. A372V consistently titers higher than wild type under all conditions tested.

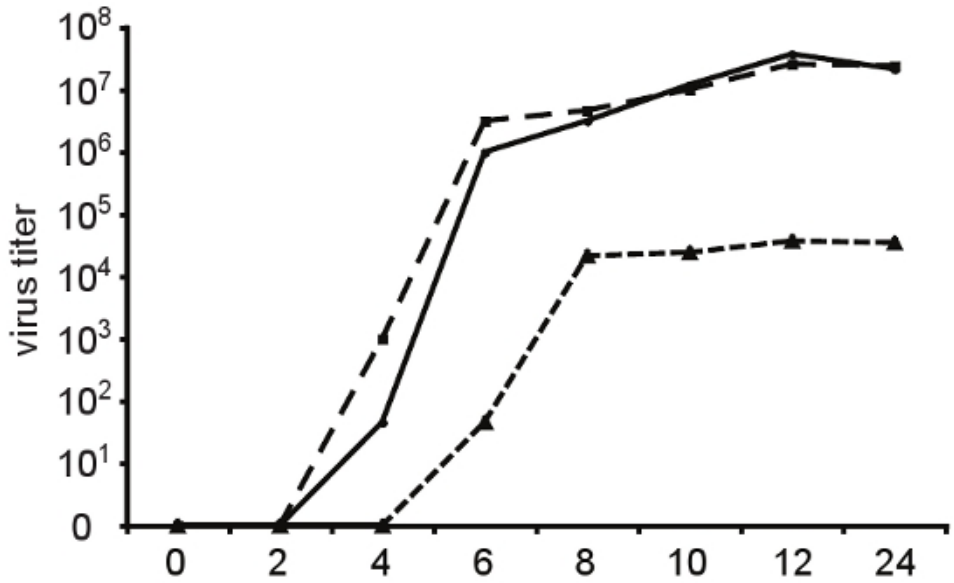

**Figure 4. Replication rates and fidelity variants.** To determine the one-step growth kinetics of virus production, HeLa cells were infected at MOI = 10 with either wild type (solid line), high fidelity variant A372V (long dashes) or replication deficient variant Cx64 (short dashes) of Coxsackie virus B3. At time points indicated, virus progeny was harvested from cells and supernatants by freeze-thaw and titered by $TCID_{50}$. The fidelity increase of A372V does not coincide with an observable replication defect in tissue culture. The variant Cx64 presents a significant delay in replication kinetics and reaches maximum titers that are 1000-fold lower than wild type virus.

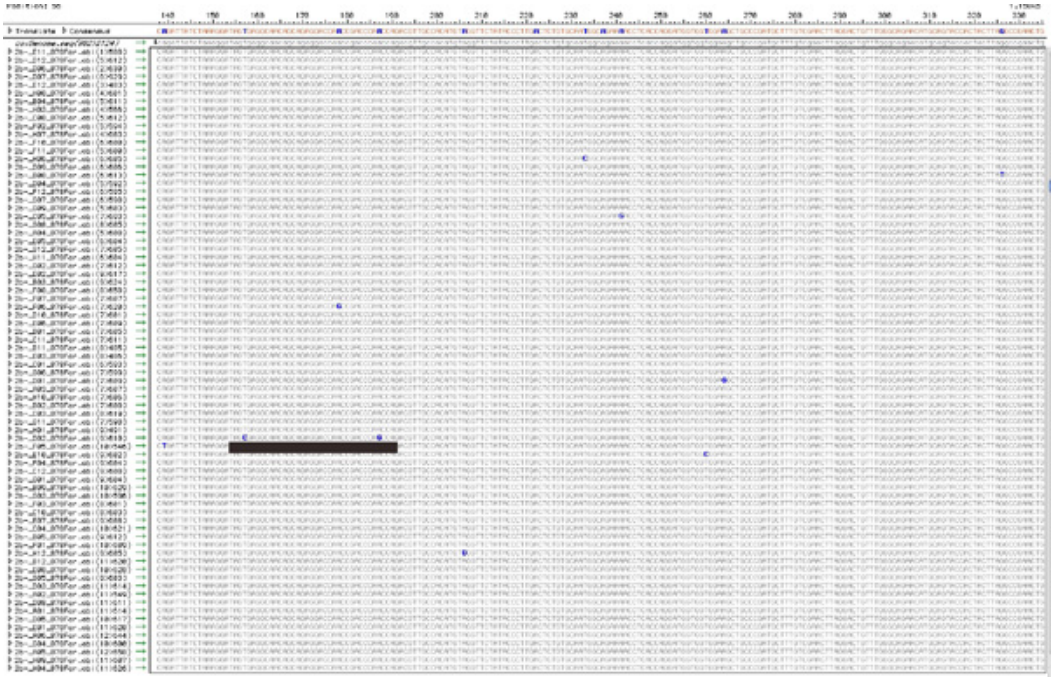

**Figure 5. Alignment of TopoTA cloned sequences from each virus population.** Using the approach described in Section 7, each sequence obtained from cloned RT-PCR product presumably originates from a single, unique genome within the total virus population and would thus, carry unique mutations. The figure shows a typical alignment, following clean up of poor quality sequences and visualization of SNPs. The total SNPs (10 in this figure) within a population are counted, and the number of SNPs appearing on each clone are noted. For example, the clone underlined by a bar, contains 2 unique mutations whereas 8 other clones contain a single, unique mutation. This data is used to compile Table 1. To view a larger version of this figure please click here.

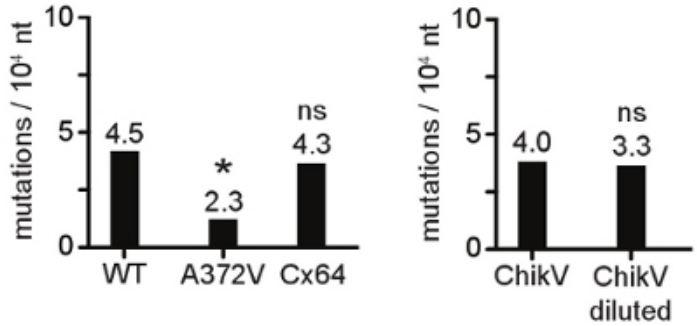


**Figure 6. Graphic representation of mutation frequencies of virus populations.** For easier interpretation, the numerical data obtained from sequence and statistical analyses can be represented as either a chart, or histogram (shown here). A372V virus generates fewer mutations than wild type and presents a significantly lower mutation frequency (*, $p<0.01$). The Cx64 variant, that replicates to titers 1000-fold lower than wild type, presents the same mutation frequency (ns, not significant) indicating that replication speed and fidelity are not necessarily linked. The same Chikungunya virus (CHIKV) population gives similar mutation frequencies whether the virus stock, or a $10^5$-fold dilution, is used for RNA extraction.

**Mutation distribution summary for statistical analysis.**

Note: For each clone, it is essential that the same genomic region (and length of sequence) is covered. In this case, 859 nucleotides per clone. This is critical for statistical analyses. On the other hand, rank sum tests used for statistical analysis do not require the sample sizes to be the same, the researcher is free to compare population of differing sample size. Hence, the 142 clones of wild type can be compared to the 84 clones of A372V.

| # clones with n mutations | wild type | A372V |
|---|---|---|
| 7 mutations | 0 | 0 |

| | | |
|---|---|---|
| 6 mutations | 0 | 0 |
| 5 mutations | 0 | 0 |
| 4 mutations | 0 | 0 |
| 3 mutations | 1 | 0 |
| 2 mutations | 6 | 2 |
| 1 mutations | 40 | 14 |
| 0 mutations | 95 | 68 |
| **Total mutations** | 55 | 18 |
| **Total clones sequenced** | 142 | 84 |
| **Total nucleotides sequenced** | 121,978 | 72,156 |
| **Mutations/$10^4$ nt** | **4.51** | **2.49** |

**Table 1. Mutation distribution summary for statistical analysis.** Note: For each clone, it is essential that the same genomic region (and length of sequence) is covered. In this case, 859 nucleotides per clone. This is critical for statistical analyses. On the other hand, rank sum tests used for statistical analysis do not require the sample sizes to be the same, the researcher is free to compare population of differing sample size. Hence, the 142 clones of wild type can be compared to the 84 clones of A372V.

## Discussion

**Choice of cell line.** The efficacy of base analogs as RNA mutagens correlates with their relative uptake by different cell types [11]. If the cell line that is normally used for virus passage proves to be refractory to mutagen uptake or too sensitive (high cell toxicity), it may be necessary to use another cell line that meets these requirements and is still permissive to viral replication. Once the mutagen resistance variant is isolated, the remainder of the characterization can be performed in the original, preferred cell line. In our experience, HeLa cells readily take up mutagen; BHK cells require up to 10-fold higher concentrations and Vero cells are refractory to mutagen uptake.

**Choice of mutagen.** In trying to isolate fidelity variants by mutagen treatment, the likelihood of success is increased if more than one type of mutagen is used. Base analog mutagens of different structure, that are erroneously incorporated into genomes during replication will predominantly induce result in a specific subset of mutations in subsequent replication cycles: ribavirin treatment favors GtoA and CtoU transition mutations [12]; 5-azacytidine has a similar bias, with the addition of CtoG and GtoC transversions [13]; 5-fluorouracil preferentially induces AtoG and UtoC transitions [14]. Alternatively, higher concentrations of $Mg^{2+}$ or $Mn^{2+}$ can be supplemented to the medium to increase the overall mutation frequency of RNA viruses without the bias described above [12]. Depending on the virus' codon sequences, and the codon changes required to generate a fidelity variant, some of these conditions will favor the emergence of this variant over others. For the higher fidelity poliovirus G64S and Coxsackie virus A372V, ribavirin treatment most readily selected for the variants because the required AtoG transition at the codon site corresponded to the mutations predominantly generated by this ribavirin.

**MOI vs. population size.** In virology, protocols for tissue culture infection pay particular attention to the multiplicity of infection (MOI), to avoid the accumulation of defective interfering particles (low MOI) or to promote recombination between viruses (high MOI), for example. To select for emergence events over serial passaging, it is also important to consider virus population size. Since the resistant mutant initially exists at low frequency, it is best to transfer as large a population size as possible from one passage to the next ($10^5$-$10^6$ viruses, e.g.) to avoid losing these emerging variants at each passage. Scaling up the size of well or flask (number of cells infected) may help to minimize the increase in MOI if this is of concern. On the other hand, for experiments in which the sensitivity of a virus to mutagen is being tested, low MOI infection is performed in order to increase the number of replication cycles occurring in the experiment and to avoid rescue of mutagenized genomes by higher fitness genomes through complementation in co-infected cells. This is important since the mutations generated on progeny genomes during the first round of replication will not be immediately detected. Most of these mutagenized RNAs will still be packaged into virion. It is in the next round of infection that lethal mutations present in these genomes will result in an aborted replication cycle, and reduction in virus titer. It may be necessary to allow for several rounds of accumulation of mutations before a significant effect of lethal mutagenesis is observed. Finally, if over the passage series in the presence of mutagen, the virus titers continue to drop until extinction, then the researcher should try passaging virus in gradually increasing amounts of mutagen (starting from a very low concentration).

**Isolation and generation of the RNA mutagen resistant clone from the RNA mutagen-resistant population.** RNA mutagens introduce multiple random mutations to each genome, but selection for resistance will only enrich (and fix to consensus sequence) the resistance mutation. To identify this mutation, we sequence the mutagen resistant population (consensus of the population) and not individual viruses. Hence, the single, random mutations created by the mutagen are not detected in the sequence; only the mutations that result in consensus changes following selection, are found. In our experience we only identify one or two such consensus sequence changes. Once the mutagen resistant population is obtained and the resistance mutation is identified, it is necessary to generate a more pure stock of this variant. Above, we described a plaque purification procedure. Alternatively, if the virus of interest does not produce easily identifiable plaques, the desired variant may be purified by limiting dilution. This approach is essentially a $TCID_{50}$ in 96-well format, where the virus stock is diluted such that less than 50% of wells are infected. Using this dilution, the same approach as above is taken, in isolating up to 10 individual variants and confirming their sequences. As mentioned, in the best cases, an infectious cDNA clone of the virus strain is available. Isolation of the variant would thus not be necessary. In our experience, fidelity variants are the result of single amino acid substitutions and can thus be generated using simple, commercialized mutagenesis kits such as Quikchange (Agilent). A secondary option is to use a cDNA clone of a closely related strain. However, if a related strain is used, we strongly suggest using both this approach and virus isolation (e.g. plaque purification) because we have found that the same fidelity altering mutation on two closely related viruses will not necessarily have the same effect.

**Fidelity and replication.** Selection of RNA mutagen resistant variants have resulted in the isolation of both higher and lower fidelity variants with growth characteristics that are similar to their wild type counterparts [4,12,15]. Currently, the link between polymerase activity rates and fidelity is not fully understood. In vitro biochemical studies using purified RNA polymerase have shown that higher fidelity variants have slower processing rates, while lower fidelity variants tend to have faster processing [1-3,12]. In tissue culture, these differences are not usually evident, suggesting that availability of resources, rather than intrinsic polymerase activity kinetics, is the rate-limiting step. If the fidelity variant replicates with kinetics that are not significantly different from wild type, then a comparison of their mutation frequencies can be directly made. If a very significant change in replication kinetics exists, then the data should be normalized to account for kinetic differences, for example by comparing viruses that have undergone the same number of replication cycles. In our experience, although no significant differences in one-step growth kinetics were observed between wild type and high fidelity variants, we observed that higher fidelity variants consistently titer higher (within 1 log) compared to wild type but they make slightly less RNA (within the same order of magnitude), further suggesting that the genomes they produce contain fewer mutations and are thus, more infectious.

**Sample preparation and sequencing.** For all steps in these protocols, it is imperative that high-fidelity, proof-reading enzymes are used for PCR and RT-PCR to limit introducing additional mutations since they cannot be distinguished from biologically relevant mutations. It is critical that the virus populations to be compared have been prepared in the same conditions (passage history, tissue culture medium, temperature, RNA extraction method, RT-PCR protocols, etc.) It is also important to ensure that enough starting material was obtained from the RNA extraction such that a strong band is generated by RT-PCR. A 1/100 dilution of the RNA sample should also give a detectable RT-PCR band, indicating that the sample contains sufficient numbers of RNA molecules to avoid representation bias (amplifying the same genome repeatedly). Since mutation frequency is a distribution, one would expect that similar values will be obtained regardless of population size, provided the aforementioned bias is not occurring. As Figure 6 shows, a $10^5$-fold dilution of a virus stock gives a mutation frequency that is not significantly different from the parental stock.

Until the optimal conditions for TopoTA cloning are found, confirm the presence of inserts after blue/white screening, by colony PCR before sequencing. As a control for mutational noise (mutations introduced by RT-PCR and sequencing), clone a PCR product from a plasmid bearing the same viral sequence and/or clone and sequence RT-PCR products of in vitro transcribed RNA corresponding to virus genome (be aware that different in vitro transcription enzymes have different error rates and may not give useful information as to the real background error in your procedure). Some virus sequences may be toxic to bacteria, so it is important to verify this before deciding on the region of viral genome to be sequenced for mutation frequency. In analyzing sequences obtained by TopoTA, note that each clone should contain only one insert/sequence. If a double peak is observed, suggesting a mixed population, it is possible that two neighboring bacterial colonies were selected. It is also possible, although highly unlikely given the low mutation frequencies in bacterial replication, that the mutation was introduced during amplification of plasmid in the bacterial culture. In plaque purified populations, a double peak may represent overlapping plaques, or a virus that is acquiring a new mutation or reverting a mutation during plaque development. Be consistent and decide on whether to count or not count these mutations.

Finally, keep in mind that the mutation frequencies used here are relative values. They are valid only in comparing virus populations grown under the same condition, and sequenced over the same region! They should not be taken as absolute values of mutation rate, or the mutation frequency of the genome as a whole. However, when conditions are controlled, they do permit reproducible, quantitative comparisons of differences in mutation distribution and frequency.

## Disclosures

No conflicts of interest declared.

## Acknowledgements

This work was supported by funding from the Medical and Health Research grant from the City of Paris, the French National grant ANR-09-JCJC-0118-1, and the ERC Starting Grant RNAvirusPopDivNVax Project no. 242719.

## References

1. Arias, A. *et al*. Determinants of RNA-dependent RNA polymerase (in)fidelity revealed by kinetic analysis of the polymerase encoded by a foot-and-mouth disease virus mutant with reduced sensitivity to ribavirin. *J Virol* **82**, 12346-12355 (2008).
2. Arnold, J. J., Vignuzzi, M., Stone, J. K., Andino, R., & Cameron, C. E. Remote site control of an active site fidelity checkpoint in a viral RNA-dependent RNA polymerase. *J Biol Chem* **280**, 25706-25716 (2005).
3. Korneeva, V. S., & Cameron, C. E. Structure-function relationships of the viral RNA-dependent RNA polymerase: fidelity, replication speed, and initiation mechanism determined by a residue in the ribose-binding pocket. *J Biol Chem* **282**, 16135-16145, doi:M610090200 [pii] 10.1074/jbc.M610090200 (2007).
4. Pfeiffer, J. K., & Kirkegaard, K. A single mutation in poliovirus RNA-dependent RNA polymerase confers resistance to mutagenic nucleotide analogs via increased fidelity. *Proc Natl Acad Sci U S A* **100**, 7289-7294, doi:10.1073/pnas.1232294100 1232294100 [pii] (2003).
5. Pfeiffer, J. K., & Kirkegaard, K. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog* **1**, e11, doi:10.1371/journal.ppat.0010011 (2005).
6. Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E., & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344-348 (2006).
7. Vignuzzi, M., Wendt, E., & Andino, R. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat Med* **14**, 154-161 (2008).
8. Crotty, S., Cameron, C., & Andino, R. Ribavirin's antiviral mechanism of action: lethal mutagenesis? *J Mol Med* **80**, 86-95, doi:10.1007/s00109-001-0308-0 (2002).
9. Coffey, L. L., & Vignuzzi, M. Host alternation of chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. *J Virol* **85**, 1025-1035, doi:JVI.01918-10 [pii] 10.1128/JVI.01918-10 (2011).

10. Ciota, A. T. *et al*. Role of the mutant spectrum in adaptation and replication of West Nile virus. *J Gen Virol* **88**, 865-874, doi:88/3/865 [pii] 10.1099/vir.0.82606-0 (2007).
11. Ibarra, K. D., & Pfeiffer, J. K. Reduced ribavirin antiviral efficacy via nucleoside transporter-mediated drug resistance. *J Virol* **83**, 4538-4547, doi:JVI.02280-08 [pii] 10.1128/JVI.02280-08 (2009).
12. Levi, L. I. et al. Fidelity variants of RNA dependent RNA polymerases uncover an indirect, mutagenic activity of amiloride compounds. *PLoS Pathog* **6**, e1001163, doi:10.1371/journal.ppat.1001163 (2010).
13. Sierra, S., Dávila, M., Lowenstein, P. R., & Domingo, E. Response of foot-and-mouth disease virus to increased mutagenesis: influence of viral load and fitness in loss of infectivity. *J Virol* **74**, 8316-8323 (2000).
14. Ruiz-Jarabo, C. M., Ly, C., Domingo, E., & de la Torre, J. C. Lethal mutagenesis of the prototypic arenavirus lymphocytic choriomeningitis virus (LCMV). *Virology* **308**, 37-47, doi:S0042682202000466 [pii] (2003).
15. Sierra, M. et al. Foot-and-mouth disease virus mutant with decreased sensitivity to ribavirin: implications for error catastrophe. *J Virol* **81**, 2012-2024 (2007).